National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

THE UNIVERSITY OF ALBERTA

Genetic Aspects of *Helicobacter pylori* and *Campylobacter jejuni*

by

©

Koji Hiratsuka

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND

RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MEDICAL MICROBIOLOGY AND IMMUNOLOGY

EDMONTON, ALBERTA

FALL 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-06225-2

Canada

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Koji Hiratsuka

TITLE OF THESIS: Genetic aspects of *Helicobacter pylori* and *Campylobacter jejuni*

DEGREE FOR WHICH THESIS WAS PRESENTED: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED: 1995

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) .....

PERMANENT ADDRESS:

18911-94 Avenue

Edmonton, Alberta

Canada T5T 4X5.....

DATED October 2....., 1995

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Genetic Aspects of *Helicobacter pylori* and *Campylobacter jejuni*** submitted by Koji Hiratsuka in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Diane E Taylor
.....
Dr. D.E. Taylor (Supervisor)

Janet A Robertson
.....
Dr. J.A. Robertson

A.M. Joffe
.....
Dr. A.M. Joffe

Brenda Leskiw
.....
Dr. B. Leskiw

Douglas E Berg
.....
Dr. D.E. Berg (External Examiner)

Date 22 Sept. 1995

This thesis is dedicated to Ursula, my partner in life.
Your dedication to our family, the sacrifices you have made
and your love have made this thesis possible.

Abstract

Genomic libraries of *Helicobacter pylori* UA802 were constructed, using the lambda cloning vector λ-DASH II. From these libraries, ribosomal RNA (rRNA) genes, and genes encoding proteins involved in copper transport (*hpCop*A and *hpCop*P) were isolated. The DNA sequences of the 23S and 5S rRNA genes were determined, and used for phylogenetic classification of *H. pylori* UA802. The phylogenetic studies showed that *H. pylori* UA802 is closely related to *Campylobacter* spp., but belongs in its own distinct genus.

Pulsed-field gel electrophoresis (PFGE) analyses were performed on *H. pylori* UA802. From these data a genomic restriction map was constructed with the enzymes *Nru*I, *Not*I, and *Sfi*I. A number of genes were localized on the map by hybridization to Southern blots of PFGE gels.

In 1990 two *Campylobacter*-specific DNA probes, pDT1719 and pDT1720, were developed. For this research, the *C. jejuni* UA580 DNA inserts of these two probes were sequenced. Two partial open reading frames (ORF), ORF1 and ORF2, were found in pDT1719 and one partial ORF, ORF3, was found in pDT1720. Computer homology searches were conducted for both the nucleotide and predicted amino acid sequences of the three ORFs. No sequences homologous to ORF3 could be found. The predicted amino acid sequence of ORF1 was found to have homology with the HtrA proteins of *S. typhimurium* and *E. coli*. The predicted amino acid sequence of ORF2 had

homology with several proteins which have been associated with the environmental response elements of two-component regulatory systems. ORF1 and ORF2 appeared to be linked together in an operon and may represent a two-component regulatory system of *C. jejuni* which responds to heat shock.

The N-terminal sequence of the first thirty amino acids of the major outer membrane protein (MOMP) of *C. jejuni* UA580, which has been shown to be a porin, was obtained. This amino acid sequence was found to have 59% identity with the amino acid sequence of two membrane proteins of *Campylobacter rectus*, one of which has also been shown to be a porin.

Acknowledgements

I would like to thank Dr. Diane Taylor for giving me the opportunity to carry out my graduate studies in her lab. I would also like to thank my committee members, Dr. A.M. Joffe, Dr. R. Read and Dr. J. Robertson, for all their assistance in the preparation of this thesis, and for their patience and guidance. Special thanks to Dr. Robertson for the extra time spent proof-reading this thesis and lighting the necessary fires; the support was greatly appreciated. I would also like to thank Dr. A.M. Joffe for agreeing to be a member of my committee on short notice and for altering his schedule to accommodate my oral examination. Thanks also to Dr. B. Leskiw and Dr. D.E. Berg for acting as my external examiners.

My research was funded by grants obtained by Dr. Taylor from Glaxo and by the Department of Medical Microbiology and Immunology.

I am grateful to my co-workers for all the help I have received over the years. It is so nice to know that I can count so many people as friends. I should acknowledge all the help I received with PCR and mapping from both Elizabeth Newnham and Qin Jiang. I would also like to give special thanks to Nick Chang, Andrew Chau, and William Yan for years of friendship and valued advice.

Many people in the Department of Medical Microbiology have influenced me over the past several years. My heartfelt thanks go out to all who have helped me.

Finally I would like to thank my family; my parents for their support over the course of my studies, and Erik, Stefan, and Tony for helping me to keep life in perspective. Most of all I would like to thank Ursula, for supporting me in so many ways; for all the late nights proof-reading this thesis, for years of love and sacrifice, and for believing in me even when I did not believe in myself.

Table of contents

Chapter 1 Introduction 1

A. *Helicobacter pylori* 1

A.1. The history of *H. pylori* 1

A.2. Microbiology and pathogenesis of *H. pylori* 2

A.3. Genetics of *H. pylori* 9

B. P-type ATPase of *H. pylori* 10

C. Ribosomal RNA 11

D. Genomic studies of *H. pylori* 13

E. The major outer membrane protein of *Campylobacter jejuni* 14

F. Objectives 15

F.1. *H. pylori* 14

F.2. *C. jejuni* MOMP 16

Chapter 2 Materials and Methods 18

A. Bacterial strains, growth conditions and vectors 18

A.1. *Helicobacter pylori* strains 18

A.2. *Campylobacter jejuni* 20

A.3. *Escherichia coli* strains 20

A.4. Vectors 22

B. Cloning techniques 22

B.1. Isolation of chromosomal DNA 22

B.2. Restriction endonuclease digestion of DNA 24

B.3. Sodium chloride gradient size fractionation of DNA 26

B.4. Preparation of vector DNA for cloning 27

B.5. Ligation of vector to passenger DNA 28

B.6. Transformation of competent cells 29

B.7. *In vitro* packaging and transfection of lambda and cosmid DNA 30

B.8. Replica filters for screening 31

B.9 Isolation of plasmid/cosmid DNA 34

B.10 Isolation of lambda DNA 37

C. DNA manipulation techniques 40
C.1. Agarose gel electrophoresis 40
C.2. Southern transfer 41
C.3. Preparation of probes for hybridization 41
C.4. Hybridization 43
C.5. Low melting point agarose extraction of DNA fragments 45
C.6. Polymerase chain reaction (PCR) methods 46
C.7. Partial digestion mapping of lambda clones 47
D. Pulsed Field Gel Electrophoresis 48
D.1. Preparation of blocks for PFGE 48
D.2. Restriction digestion of embedded DNA 49
D.3. PFGE gels 49
E. DNA sequencing 50
E.1. Double-stranded dideoxy-sequencing 50
E.2. Cycle sequencing with ^{33}P 50
E.3. Sequence analysis 50
E.4. Construction of deletion mutants for sequencing 52
F. RNA techniques 53
F.1. Isolation of total RNA 53
F.2. Denaturing RNA agarose gels 54
F.3. Primer extension 54
G. SDS-PAGE and western blots 55

Chapter 3 Construction of genomic libraries of *H. pylori* 57
A. Introduction 57
A.1. Plasmid vectors 57
A.2. Cosmid vectors 59
A.3. Lambda vectors 60
B. Materials and Methods 61
C. Results 61
C.1. Construction of plasmid libraries of *H. pylori* UA763 61
C.2. Construction of cosmid libraries of *H. pylori*

DNA 64
C.3. Construction of lambda genomic libraries of *H. pylori* DNA 66
D. Discussion 69
D.1. Plasmid libraries 69
D.2. Cosmid library construction 74
D.3. Lambda library construction 76

Chapter 4 Isolation and characterization of the P-type ATPase and rRNA genes from the genomic libraries of *H. pylori* strain UA802 ... 79
A. Introduction 79
A.1. P-type ATPase of *H. pylori* 79
A.2. The 23S and 5S genes of *H. pylori* UA802 79
B. Materials and Methods 80
C. Results 80
C.1. Cloning of a P-type ATPase of *H. pylori* UA802 ... 80
C.2. PCR amplification of rRNA probes and cloning of the rRNA genes of *H. pylori* UA802 85
D. Discussion 114
D.1. Evidence for a P-type ATPase in *H. pylori* UA802 and UA763 114
D.2. Cloning and characterization of the rRNA genes of *H. pylori* UA802 118

Chapter 5 Pulsed-field gel electrophoresis analysis and revision of the genomic map of *H. pylori* strain UA802 128
A. Introduction 130
B. Materials and Methods 130
C. Results 130
C.1. Construction of the *H. pylori* UA802 genomic restriction map 130
D. Discussion 143
D.1. PFGE analysis and construction of a genomic

restriction map of *H. pylori* UA802 143
D.2. Comparisons of genome arrangements 145

Chapter 6 *Campylobacter jejuni* outer membrane proteins 149
A. Introduction 149
B. Materials and Methods 150
C. Results 150
C.1. DNA sequence analysis of pDT1719 and pDT1720 150
C.2. N-terminal amino acid sequencing of the MOMP of *C. jejuni* UA580 161
D. Discussion 166
D.1. Characterization of pDT1719 and pDT1720 166
D.2. Comparison of the N-terminal sequence of the *C. jejuni* MOMP to the N-terminal sequence of the porin protein of *Campylobacter rectus* 168

Chapter 7 General Discussion 170
A. *Helicobacter pylori* genetics 170
A.1. Genomic libraries of *H. pylori* UA802 170
A.2. Cloning and characterization of the *H. pylori* UA802 *hpCop*A, *hpCop*P, and rRNA genes 173
A.3. Genomic restriction map and PFGE analyses of *H. pylori* UA802 179
B. Analysis of the *C.jejuni* DNA probes and MOMP 181
B.1. Characterization of the *C.jejuni* DNA probes 181
C. Future study 182

Bibliography 185

Appendix A. Composition of media and solutions 207

Appendix B. PCR primers used to amplify gene probes from *H. pylori* UA802 210

Appendix C. Comparisons of rRNA gene sequences from *Bacillus subtilis*..... 211

List of Tables

Table 1-1. Members of the genus *Helicobacter* 3

Table 2-1. *Helicobacter pylori* strains used in this research 19

Table 2-2. *Escherichia coli* strains used in this research 21

Table 3-1. Transformation efficiency of plasmid DNA in *E. coli* DH5α and DH5αMCR strains 63

Table 3-2. Efficiency of cosmid libraries in *E. coli* LE392 65

Table 3.3. Transfection efficiency of lambda libraries ... 67

Table 3-4. Restriction phenotypes of some *E. coli* strains 70

Table 4-1. Primers used for PCR amplification of 16S and 23S rRNA probes 88

Table 4-2. Primers used for sequencing and primer extension of the 23S and 5S genes of *H. pylori* UA802 99

Table 5-1. Restriction fragment sizes of *H. pylori* UA802 DNA 131

Table 5-2. *H. pylori* UA802 fragment localization 133

Table 5-3. Probes used for localization of genes in the *H. pylori* UA802 genome 138

Table 5-4. Localization of genes to the *H. pylori* UA802 142

List of Figures

Figure 4-1. Hybridization of restriction digestions of *H. pylori* strains UA763 and UA802 with cloned P-type ATPases 82

Figure 4-2. Regions of conserved amino acid sequence in P-type ATPases used for designing of PCR primers 84

Figure 4-3. Agarose gel analysis of P-type ATPase PCR probe amplified from *H. pylori* UA802 86

Figure 4-4. Restriction digestion analysis and hybridization of λP9 and λP14 87

Figure 4-5. Agarose gel analyses of rRNA probes amplified from *H. pylori* UA802 using PCR 89

Figure 4-6. Agarose gel electrophoresis of *Eco*RI and *Xba*I restriction digestions of clones λ16.4 and λ23.54, and hybridization of Southern transfers of the gels with either the 16S or 23S PCR probes 91

Figure 4-7. Low resolution restriction maps of clones λ16.4 and λ23.54 constructed using partial digestion mapping, showing the location of the 16S and 23S genes 93

Figure 4-8. Agarose gel electrophoresis of *Xba*I digestions of nine 23S λ-clones and hybridization of these clones with DNA from the region flanking the 23S gene of λ23.54 ... 95

Figure 4-9. Restriction maps of p54X1.1 and pB1X1.1 subcloned from λ23.54 and λ23.B1 respectively, showing the location of the 23S and 5S genes and sequencing strategy used to determine rRNA gene sequences 97

Figure 4-10. Complete nucleotide sequences of the 23S and 5S genes as well as part of the flanking region from clones p54X1.1 and

pB1X1.1 100
Figure 4-11. Primer extension of total RNA of *H. pylori* UA802 to determine the start points of the 23S and 5S rRNA molecules 105
Figure 4-12. Sequence of DNA from the points where p54X1.1 and pB1X1.1 diverge 108
Figure 4-13. PCR amplification from p54X1.1, pB1X1.1, λ23.54, λB1 and *H. pylori* UA802 chromosomal DNA using primers designed from sequences 3' to the 5S gene where the p54X1.1 and pB1X1.1 sequences differ 110
Figure 4-14. Phylogenetic tree based on comparisons of complete 23S rRNA gene sequences, including representative organisms from four of the five Proteobacteria subdivisions and representative organisms from the Gram-positive (*B. subtilis* and *M. luteus*) and cyanobacteria (*A. nidulans*) phyla 112
Figure 4-15. Phylogenetic tree based on comparisons of available partial 23S rRNA gene sequences of *Campylobacter* species and *H. pylori* 113
Figure 4-16. Phylogenetic tree based on comparisons of complete 5S rRNA gene sequences from representative organisms of eight of the ten Eubacterial phyla 115

Figure 5-1. Restriction map of *H. pylori* UA802 constructed by Taylor et al., 1992a 129
Figure 5-2. Fragment hybridizations of PFGE Southern blots with PFGE restriction digestion fragments isolated using LMP-agarose 135
Figure 5-3. Revised genomic restriction map of *H. pylori* UA802 137
Figure 5-4. Hybridization of Southern blots of PFGE gels of *H. pylori* UA802 DNA with various

gene probes 140
Figure 5-5. Alignment of genomic maps of *H. pylori* strains NCTC11638, NCTC11639, UA861 and UA802 144

Figure 6-1. Restriction maps of pDT1719 and pDT1720, and sequencing strategy used to determine the sequence of the cloned DNA 151
Figure 6-2. The complete nucleotide sequence and predicted amino acid sequences of the *C.jejuni* DNA cloned in the *Campylobacter* DNA probe, pDT1719 153
Figure 6-3. The complete nucleotide sequence and predicted amino acid sequence of the *C.jejuni* DNA cloned in the *Campylobacter* probe, pDT1720 156
Figure 6-4. Amino acid sequence comparisons of the *C. jejuni* HtrA protein with sequences found in Genbank 159
Figure 6-5. Amino acid sequence comparisons of the *C. jejuni* HtrB protein with sequences found in Genbank 162
Figure 6-6. SDS-PAGE of purified MOMP samples from *C. jejuni* UA580 164
Figure 6-7. N-terminal amino acid sequence of the *C. jejuni* UA580 45 kDa MOMP and comparison with the N-terminal sequence of the *C. rectus* 43 kDa and 51 kDa membrane proteins 165

Figure 7-1. Comparison of the domain V loops of the 23S rRNA molecules from *H. pylori* UA802, *Bacillus subtilis* and *Mycobacterium intracellulare* 177

List of Abbreviations

aa	amino acid
BHI	Brain Heart Infusion
BMC	Boerhinger Mannheim Canada
bp	base pair
Brinkman	Brinkman Instruments (Canada) Ltd.
BRL	Gibco/BRL Life Technologies Inc.
BSA	bovine serum albumin
CAPS	3-[cyclohexylamino]-1-propanesulfonic acid
CFU	colony forming units
Da	Daltons
DEPC	diethyl pyrocarbonate
dH_2O	Distilled water
Difco	Difco Laboratories
DIG	digoxygenin
DNase	deoxyribonuclease
EDTA	ethylenediamine tetra acetic acid
ETOH	ethanol
GCG	Genetics Computer Group
GM	gastric metaplasia
hr	hour
ICN	ICN Biochemicals Canada Ltd.
kb	kilobase
kDa	kilodalton
LB	Luria broth
LBM	Luria broth + 0.5% maltose + 10mM $MgSO_4$

LMP	low melting point
MeOH	methanol
min	minute
MOMP	major outer membrane protein
MOPS	morpholinopropanesulfonic acid
NaOAc	sodium acetate
NT	nick translation
nt	nucleotide
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	polyethylene glycol
PFGE	pulsed-field gel electrophoresis
PFU	plaque forming units
Pharmacia	Pharmacia Biotech Inc.
PMSF	phenylmethylsulfonyl fluoride
RNase	ribonuclease
rRNA	ribosomal RNA
r·DNA	ribosomal DNA
s	second
SDS	sodium dodecyl sulfate
Sigma	Sigma Chemical Company
SSC	salt/sodium citrate
TAFE	Transverse Alternating Field Electrophoresis
TB	Tris borate

TE	10 mM Tris·HCl pH 7.5 1 mM EDTA
TEMED	N,N,N',N'-Tetra-methylethylenediamine
UAR	upstream activation region
USB	United States Biochemica

Chapter 1
Introduction

A. *Helicobacter pylori*

Studies of the bacterium *Helicobacter pylori* have changed our view of human gastric disease. The human stomach has not normally been considered a hospitable environment for bacterial growth, so suggestions that *H. pylori* colonized the stomach and caused gastroduodenal disease were initially met with a great deal of skepticism (Marshall, 1983; Warren, 1983). Research over the past several years has gradually removed this doubt, and *H. pylori* has now become widely accepted as an etiological agent for chronic type B gastritis, characterized by gastric inflammation (Marshall, *et al.*, 1985; Goodwin *et al.*, 1986; Cover & Blaser, 1992a). *H. pylori* infection is also believed to play a role in both peptic and duodenal ulceration (Goodwin *et al.*, 1986, Dixon, 1991). In addition, there is evidence that chronic infection with *H. pylori* increases the risk of gastric adenocarcinoma (Hebbel, 1940; Correa *et al.*, 1990; Forman *et al.*, 1990; Talley *et al.*, 1991; Parsonnet *et al.*, 1991) and gastric lymphoma (Parsonnet *et al.*, 1994).

A.1. The history of *H. pylori*

Although spiral Gram-negative organisms had been observed in human gastric biopsies as early as 1874, the involvement of these organisms in the pathogenesis of gastroduodenal

disease was unclear. Culturing these organisms *in vitro* proved difficult. Thus, aside from microscopic observations, very little was known about them. In 1983 the organism that we now call *Helicobacter pylori* was first cultured from human gastric biopsy specimens in Perth, Australia (Marshall, 1983; Warren, 1983). Initially the name *Campylobacter pyloridis* was chosen because of physiological and morphological similarities of *C. pyloridis* with organisms from the genus *Campylobacter* (Marshall *et al.*, 1987). The name was later changed to *C. pylori* (Marshall and Goodwin, 1987). However, even early observations of flagellar morphology suggested that there were obvious differences which set this organism apart from *Campylobacter* spp. (Marshall and Warren, 1984). Later studies of fatty acid profiles, ultrastructure (Goodwin *et al.*, 1987) and ribosomal RNA (Romaniuk *et al.*, 1987) suggested that the organism designated *C. pylori* was not a true *Campylobacter* and that the creation of a new genus was in order. In 1989 the genus *Helicobacter* was created and *H. pylori* was included in this genus along with the animal pathogen *H. mustelae* (Goodwin *et al.*, 1989). Currently there are twelve species in the genus *Helicobacter* (Table 1-1), but with improved methods for the isolation and cultivation of organisms, this number could increase.

A.2. Microbiology and pathogenesis of *H. pylori*

A number of reviews have been published describing the morphological and physiological characteristics of *H. pylori*

Table 1-1. Members of the genus *Helicobacter*[a]

Species	Host	Source
H. pylori	Human, monkey, cat	Stomach
H. nemestrinae	Macaque monkey	Stomach
H. acinonyx	Cheetah	Stomach
H. heilmannii	Mammals	Stomach
H. felis	Cat, dog	Stomach
H. rappini	Mammals	Stomach, intestine
H. mustelae	Ferret	Stomach
H. muridarum	Rat, mouse	Stomach, intestine
H. hepaticus	Mouse	Liver
H. canis	Dog	Intestine
H. fenneliae	Human	Intestine
H. cinaedi	Human	Intestine

a- From Cover and Blaser, 1995

(Dick, 1990; Good and Armstrong, 1990; Goodwin and Worsley, 1993). *H. pylori* is a Gram-negative organism which exists as either a curved rod when grown on solid media or as a spiral rod in liquid media or *in vivo*. It is highly motile, with four to six polar, sheathed flagella. It is a microaerophilic organism and, upon prolonged exposure to oxygen, will assume a viable but non-culturable coccoid form. *H. pylori* has a low G + C content calculated to be 35.2 mol% (Béji *et al.*, 1988). The genome size of *H. pylori* has been estimated to be approximately 1700 kb (Taylor *et al.*, 1992a; Bukanov and Berg, 1994; Jiang, 1994).

Enzymology

A variety of enzymes are produced by *H. pylori*, including large amounts of urease (Owen *et al.*, 1986; Mobley *et al.*, 1988), superoxide dismutase and catalase (Lior & Johnson, 1985), as well as proteases (Slomiany *et al.*, 1987; Baxter *et al.*, 1989), lipase (Sarosiek *et al.*, 1988), DNase, and oxidase. Urease is of particular interest, as the production of this extracellular enzyme distinguishes *H. pylori* from other spiral organisms. Urease activity has been used as the basis of diagnostic tests to detect the presence of *H. pylori* (McNulty & Wise, 1985; Marshall & Surveyor, 1988). In addition, numerous studies have indicated that urease is one of the important factors in colonization of the stomach by *H. pylori* (Hazell and Lee, 1986; Marshall. *et al.*, 1990; Eaton *et al.*, 1991; Triebling *et al.*, 1991; Segal *et al.*, 1992; Mégraud *et al.*, 1992). It has been hypothesized that urease

activity produces a cloud of ammonia protecting the bacterium from the acidic environment of the stomach long enough for the bacterium to penetrate the mucosal lining. The two structural genes *ure*A and *ure*B encode the urease subunits and were among the first genes from *H. pylori* to be characterized (Clayton *et al.*, 1989a, 1989b; Labigne *et al.*, 1991; Cussac *et al.*, 1992). To date nine genes are thought to be involved in the expression of urease activity in *H. pylori* (Cover and Blaser, 1995).

Flagella

The flagella have also been implicated in the pathogenesis of *H. pylori*. Four to six uni-polar flagella are normally present, although bipolar flagella have been observed under scanning electron microscopy. The sheath sometimes extends to form a terminal bulb at the end of each flagellum. The flagella confer motility to the organism and, combined with the spiral shape of the organism, are thought to aid in the penetration of the mucous layer and the colonization of the stomach lining (Eaton *et al.*, 1989). *H. pylori* has two flagellin genes, designated *fla*A (Leying *et al.*,1991) and *fla*B (Suerbaum *et al.*, 1993). Studies have shown that *fla*A, but not *fla*B, is required to assemble functional flagella (Haas *et al.*, 1993; Suerbaum *et al.*, 1993).

Adhesin

H. pylori specifically colonizes the mucosal epithelium of the stomach with a marked preference for the antrum (Blaser, 1987). The bacteria are found closely associated with

mucosal cells and can be observed in the intercellular junctions (Hazell *et al.*, 1986, Chen *et al.*, 1986). As well, *H. pylori* cells have been observed colonizing the duodenum at sites of gastric metaplasia (GM)(Wyatt *et al.*, 1987; Wyatt *et al.*, 1990). This observed tissue-specificity suggests the existence of *H. pylori*-specific receptors in the antrum of the stomach and in GM tissue in the duodenum (Lingwood, 1993). An *N*-acetylneuraminyl-lactose-binding hemagglutinin (NLBH) associated with the cell-surface has been identified, which can specifically bind to the *N*-acetylneuraminyl-α(2-3)-galactopyranosyl (NeuAc(2-3)Gal) component of *N*-acetylneuraminyl-lactose (NeuAc-lactose) on the surface of erythrocytes (Evans *et al.*, 1988). Evans *et al.* postulated that this protein mediated the interaction between *H. pylori* and the gastric epithelium. The gene encoding this adhesin, designated *hpa*A, has been cloned and sequenced (Evans *et al.*, 1993).

The adhesion of attaching and effacing enteropathogenic *Escherichia coli* to the surface of intestinal epithelial cells elicits the formation of adhesion pedestals (Moon, *et al*, 1988). These pedestals are the result of the polymerization of F-actin. The adhesion of *H. pylori* to epithelial cells was originally characterized as a similar "attaching and effacing" interaction (Hemalatha *et al.*, 1991). However, polymerized F-actin adhesion pedestals could not be observed either *in vitro*, in cell cultures infected with *H. pylori*, or *in vivo*, in human gastric mucosal biopsies

of individuals infected with *H. pylori* (Dytoc *et al.*,1993). It has been demonstrated that phosphatidylethanolamine (Dytoc *et al.*,1993) and the human blood group antigen Lewis^b (Borén *et al.*, 1993) are important for the adhesion of *H. pylori* to human epithelial cells.

Cytotoxin

Studies of cytotoxin production by *H. pylori* revealed that about half of the strains tested produced a protein, found in the supernatants of broth cultures, which caused a non-lethal vacuolization of eucaryotic cells (Leunk *et al.*, 1988; Cover *et al.*, 1990). This cytotoxic activity was both heat-labile and protease sensitive. Two proteins of 87 kDa and 128 kDa were observed in the supernatants with cytotoxic activity.

The 87 kDa protein was shown to have the vacuolating activity (Cover & Blaser, 1992b). The gene encoding the toxin, *vac*A, was sequenced from a number of toxin-producing *H. pylori* strains (Phadnis *et al.*, 1994; Schmitt and Haas, 1994; Telford *et al.*, 1994; Cover *et al.*, 1994). Partial sequence of the *vac*A gene from a toxin-negative strain was also obtained. This sequence was significantly different when compared to the sequence of the toxin-positive *vac*A gene (Cover *et al.*, 1994). The exact role of cytotoxin in the pathogenesis of *H. pylori* is undetermined. Purified cytotoxin caused gastric lesions in mice similar to those observed in human subjects infected with *H. pylori* (Telford *et al.*, 1994). This implied that expression of the cytotoxin may play a central role in the pathogenesis of

H. pylori. Other studies indicated that patients infected with a strain of *H. pylori* expressing the cytotoxin had a higher risk of developing duodenal ulceration (Figura *et al.*, 1989; Tee *et al.*, 1993). All strains of *H. pylori* appeared to possess a *vac*A gene. However, there were marked differences in both the nucleotide and amino acid sequences between the *vac*A genes from the strain which did express cytotoxic activity and the strain which did not (Cover *et al.*, 1994).

The role of the 128 kDa protein, designated CagA (Cytotoxin Associated Gene), in expression of cytotoxic activity is not clear. Expression of the CagA protein appears to be closely associated with expression of cytotoxin (Xiang *et al.*, 1995), although the expression of cytotoxin is not dependent on CagA expression (Tummuru *et al.*, 1994). It was also found, in hybridization studies, that not all *H. pylori* strains possess a *cag*A gene. Furthermore, not all strains which possess a *cag*A gene express the CagA protein (Xiang *et al.*, 1995). The region upstream of the *cag*A gene may also be associated with cytotoxin expression, as some strains with the CagA$^-$VacA$^-$ phenotype which possessed a *cag*A gene had deletions of several kilobases (kb) in this region. The *cag*A genes from two strains of *H. pylori* have been cloned and sequenced (Tummuru *et al.*, 1993; Covacci *et al.*, 1993).

Infection with *H. pylori* has been shown to induce the production of the interleukin IL-8 as part of the inflammatory response (Crabtree *et al.*, 1994; Sharma *et al.*,

1995). Higher levels of IL-8 were induced by the CagA$^+$VacA$^+$ phenotype. However, isogenic mutants with CagA$^-$VacA$^+$, CagA$^+$VacA$^-$, or CagA$^-$VacA$^-$ phenotypes showed no decrease in IL-8 induction (Sharma *et al.*, 1995). This indicated that these two proteins were not directly involved in the IL-8 response.

A.3. Genetics of *H. pylori*

Although interest in genetic studies of *H. pylori* has increased steadily, relatively few *H. pylori* genes have been characterized. This is due, in part, to difficulties in cloning and characterizing genes from *H. pylori* (Clayton *et al.*, 1989a; Covacci *et al.*, 1993). The susceptibility of cloned *H. pylori* DNA to the host restriction modification systems, due to the methylation patterns, is thought to be one of the contributing factors to these difficulties (Phadnis *et al.*, 1993).

Aside from the genes mentioned above (the urease genes, *fla*A, *fla*B, *hpa*A, *vac*A, and *cag*A), a number of other genes have been cloned and characterized. These include a species-specific 26 kDa protein gene (O'Toole *et al.*, 1991), a bacterial ferritin gene (*pfr*) (Frazier *et al.*, 1993), a lipoprotein gene (*lpp*20) (Kostrzynska *et al.*, 1994), a superoxide dismutase gene (SOD) (Spiegelhalder *et al.*, 1993), two heat-shock protein genes (*hsp*A, *hsp*B) (Suerbaum *et al.*, 1994), a mucinase gene (*hap*) (Smith *et al.*, 1994), a gyrase gene (*gyr*A) (Moore *et al.*, 1995), a P-type ATPase gene (*hpcop*A) (Ge *et al.*, 1995) and a cation-binding protein

(*hpcopP*) (Ge *et al.*, 1995). In addition, the complete nucleotide sequences for the 16S ribosomal RNA genes of five different strains of *H. pylori* have been determined (Eckloff *et al.*, 1994).

B. P-type ATPase of *H. pylori*

Chronic infection with *H. pylori* has been associated with the development of peptic and duodenal ulceration. Treatment of ulcers has involved the use of substituted benzimidazoles such as omeprazole and lansoprazole. These compounds act by blocking sulfhydral-groups of gastric P-type ATPases thus reducing the acidity of the gastric environment and reducing the severity of ulceration. It has been found that these compounds also have anti-bacterial activity against *H. pylori* (Iwahi *et al.*, 1991; Loo *et al.*, 1992). In fact, treatment with omeprazole alone (Stolte and Bethke, 1990) or in combination with antibiotics (Unge *et al.*, 1989: Gellini *et al.*, 1991) has been shown to eradicate *H. pylori* from some infected patients.

The target site of the omeprazole in *H. pylori* has not been elucidated. Omeprazole was shown to inhibit activity of purified urease (Bugnoli *et al.*, 1993). Since urease has been shown to be an important colonization factor of *H. pylori*, this suggested that omeprazole could reduce the ability of *H. pylori* to colonize an individual by inhibiting urease activity. However, minimum inhibitory concentration

(MIC) studies of both urease positive and negative strains of *H. pylori* showed that both strains were equally inhibited by omeprazole (Bugnoli *et al.*, 1993) and that inhibition of growth by omeprazole was directed through a target site other than the urease. Although a P-type ATPase from *H. pylori* has been cloned and characterized (Ge *et al.*, 1995; this thesis) this particular ATPase did not appear to act as a target for omeprazole *in vitro*. This does not rule out the possibility of another P-type ATPase or the existence of another protein in *H. pylori* that is targeted by omeprazole.

C. Ribosomal RNA

The assembly of amino acids into proteins is a complex process involving many different components. Ribosomal RNA (rRNA) plays a central role in this process. The structure and function of rRNA is highly conserved throughout all life-forms. The similarities and differences among these molecules are used in cataloging and sorting living organisms into their appropriate phylogenetic categories (Woese, 1987). All living organisms can be placed into one of three primary kingdoms based on 16S rRNA sequence. Eucaryotes form one kingdom, while procaryotes are divided into two primary kingdoms, Eubacteria and Archaebacteria (Woese, 1987). Each kingdom has many divisions. The kingdom Eubacteria is divided into 10 major classes. The genus *Helicobacter* is found in the class *Proteobacteria*, the purple bacteria. This class was formerly divided into 4 subdivisions, alpha, beta,

gamma and delta. A fifth subdivision, epsilon, has been created based on partial 16S and 23S hybridization studies (Vandamme *et al.*, 1991). Within this subdivision there are three distinct rRNA clusters. *Helicobacter* spp. are grouped with *Wolinella succinogenes*, *Flexispira rappini*, and some yet unnamed *Campylobacter* spp. in rRNA cluster III. The majority of *Campylobacter* spp. are grouped together in rRNA cluster I, while members of the genus *Arcobacter*, along with some *Campylobacter*-like organisms, make up rRNA cluster II.

Procaryotic ribosomes contain three species of rRNA, 23S, 16S and 5S, which interact with ribosomal proteins to form the mature ribosome. Normally, single copies of the 16S and 23S genes and one or two copies of the 5S gene are clustered together in an operon. Several copies of these operons are usually found scattered throughout the genome. The genes for the rRNA species are not always linked together. This is the case in some strains of *Borrelia* spp. (Schwartz *et al.*, 1992), *Campylobacter jejuni* and *Campylobacter coli* (Taylor *et al.*, 1992b), *Mycoplasma hyopneumoniae* (Taschke *et al.*, 1986), *Mycoplasma gallisepticum* (Chen and Finch, 1989), *Leptospira interrogans* (Fukunaga and Mifuchi, 1989), *Pirellula marina* (Liesack and Stackebrandt, 1989), *Thermoplasma acidophilum* (Ree and Zimmermann, 1990), *Thermus thermophilus* (Hartmann and Erdmann, 1989) and *H. pylori* (Taylor *et al.*, 1992a; Bukanov and Berg, 1994). The fact that rRNA genes can be found separate from each other raises questions about regulation of rRNA expression. It was assumed that the

relative amounts of the various rRNA species were equal because they were transcribed together in operons (Nomura *et al.*, 1984). In *E. coli*, for example, the 16S, 23S and 5S genes are co-transcribed as a 30S rRNA precursor, which is later processed to give equimolar amounts of the three rRNA species. In organisms where the rRNA genes are not linked, some form of regulation to ensure that the rRNA species are represented in the proper proportions may be required.

D. Genomic studies of *H. pylori*

A genomic restriction map of *H. pylori* strain UA802 has been constructed using pulsed-field gel electrophoresis (PFGE) (Taylor *et al.*, 1992a). This map was made using the restriction endonucleases *Not*I and *Nru*I. A variety of genetic elements have been localized on this map. In 1994, Bukanov and Berg, through the use of a contiguous cosmid library, constructed a physical map of *H. pylori* strain NCTC11638. Comparisons of the maps produced by these two groups, as well as comparisons of PFGE patterns of different strains of *H. pylori* (Taylor *et al.*, 1992a), show a large degree of variability in both the restriction patterns and the arrangement of genetic elements from strain to strain. This genomic diversity has also been observed in restriction patterns using conventional electrophoresis (Majewski *et al.*, 1988; Clayton *et al.*, 1989), as well as arbitrary primer PCR (Akopyantz *et al.*, 1992). The reasons for this variability are unclear. A number of hypotheses have been put forth

including differences in methylation patterns, a high frequency of silent mutations (i.e. mutations which do not cause phenotypic change), genetic uptake or exchange and recombination with other strains or bacteria, as well as both global and local rearrangements of the chromosome. Environmental pressures during the course of infection of a new host may also trigger chromosomal rearrangements (Taylor *et al.*, 1992a).

E. The major outer membrane protein of *Campylobacter jejuni*

Campylobacter spp. are recognized as one of the leading causes of bacterial gastroenteritis world-wide. *C. jejuni* is the species most commonly isolated from man. In the United States alone, *C. jejuni* has been estimated to be responsible for up to two million cases of acute diarrhea each year (Blaser and Reller, 1981). The pathogenic mechanisms of *C. jejuni* in the etiology of disease are still not clear. Adhesion to the intestinal epithelium is one of the first steps in the pathogenesis of this organism. A number of outer membrane proteins have been implicated which may be involved in adhesion. One of these proteins is the 45 kDa major outer membrane protein (MOMP) which has been shown to be a porin (Huyer *et al.*, 1986).

In 1990, our laboratory developed non-radioactive DNA probes for the identification of *C. jejuni* and *C. coli* (Taylor and Hiratsuka, 1990). Initially, these probes were

isolated from a λ-gt11 expression library of *C. jejuni* UA580 using polyclonal antisera directed against the 45 kDa MOMP. The antiserum used was also reactive with a number of other minor components. Cloned DNA inserts from two positive λ-gt11 clones were further subcloned into pUC13 and these plasmid clones were used as probes. Plasmid pDT1719 contained an 1845 bp insert and could be used to identify both *C. jejuni* and *C. coli* by DNA hybridization. Under reduced stringency pDT1719 was also able to weakly hybridize *C. laridis*. The plasmid pDT1720 contained a 1475 bp insert and specifically hybridized to *C. jejuni* chromosomal DNA. Further work to characterize the cloned DNA has yet to be completed.

F. Objectives

F.1. *H. pylori*

An objective of this research was to construct genomic libraries of *H. pylori* strains and use these libraries to clone various genetic elements. These genomic libraries of *H. pylori* would provide a source of DNA for future research in isolation and characterization of other *H. pylori* genes. Construction of the genomic libraries is described in Chapter 3.

The genomic libraries were used to clone genes involved in P-type ATPase activity. The cloning of these genes is described in Chapter 4. Characterization of P-type ATPase

genes cloned in this work has been published (Ge *et al.*, 1995).

Another objective was to clone and sequence the 23S and 5S genes of *H. pylori* (Chapter 4), and to use these sequences to investigate phylogenetic relationships of *H. pylori* to other organisms. As mentioned previously, the sequences of the 16S genes from five different strains of *H. pylori* have already been deduced (Eckloff *et al.*, 1994). Studies utilizing the 23S and 5S genes of *H. pylori* have been limited to DNA-rRNA hybridizations. Complete sequences for these genes have not been reported. The nucleotide sequences of these genes will give us a clearer picture of where *H. pylori* fits phylogenetically in relation to other closely related organisms, such as *Campylobacter* species. *H. pylori* has already been associated phylogenetically with these organisms based on 16S and 23S hybridization studies.

The revision of a genomic restriction map, constructed previously by Taylor *et al.*, 1992a, was another objective of this research. These revisions included the addition of sites for a third restriction enzyme, *Sfi*I, and confirmation that *H. pylori* UA802 has two copies of rRNA genes, as observed in other strains of *H. pylori* (Bukanov and Berg, 1994; Jiang, 1994).

F.2. *Campylobacter jejuni* major outer membrane protein

The objectives for the study of the *C. jejuni* major outer membrane protein (MOMP) were two-fold. First, the characterization, including sequence analysis, of two

Campylobacter DNA probes initially isolated using MOMP specific antiserum was carried out. Sequence analysis was useful in identifying genetic elements which had been cloned. The second objective was to obtain amino-acid sequence of purified MOMP protein. No sequence data for either the nucleotide sequence or amino-acid sequence has yet been published for the MOMP of *C. jejuni*.

Chapter 2

Materials and Methods

A. Bacterial strains, growth conditions and vectors

A.1. *Helicobacter pylori* strains

The *Helicobacter pylori* strains used for this research are listed in Table 2-1. A number of different growth media support the growth of *H. pylori* (Goodwin & Armstrong, 1990, Goodwin & Worsley, 1993). For this research *H. pylori* cultures were grown on brain heart infusion (BHI) agar (Oxoid, Oxoid Ltd., Basingstoke, UK) with yeast extract (5%; Difco, Difco Laboratories, Detroit, MI), vancomycin (6 mg/l; Sigma, Sigma Chemical Company, St. Louis, MO, USA), amphotericin B (10 mg/l: Sigma) and 7.5% animal serum (equine or bovine; Morse's Biological Supplies, Edmonton, AB, Canada) as previously described (Taylor *et al.*, 1992). Cultures were incubated for 2-4 days in anaerobic jars with 5% H_2, 5% CO_2, and 85% N_2 (Linde Canada Inc., Edmonton, AB, Canada) at 37°C.

BHI broth (Oxoid) with 20% glycerol (BDH, BDH Inc., Toronto, ON, Canada) was used as cryogenic storage media for long-term storage of *H. pylori* cultures. For cryogenic storage, 1.5 ml of the BHI/glycerol medium were inoculated with a heavy loopful of *H. pylori* from a plate culture. Aliquots of 100 µl were dispensed into 1.0 ml cryogenic storage vials (Nalgene Co., Rochester, NY, USA) and the

Table 2-1. *Helicobacter pylori* strains used in this research

Strain	Source	Origin of cultures
UA802	Clinical isolate	University of Alberta Hospital, Edmonton, AB, Canada
UA763	National Collection of Type Cultures, NCTC11639	Laboratory Center for Disease Control, Ottawa, ON, Canada (LCDC)
UA764	Clinical isolate	LCDC
UA765	Clinical isolate	LCDC

cultures were flash frozen in liquid nitrogen. The frozen samples were stored at -82°C and have been maintained for at least 2 years without significant loss of viability.

A.2. *Campylobacter jejuni*

Campylobacter jejuni UA580 (NCTC11168) was used for study of the *C. jejuni* major outer membrane protein (MOMP) and the development of a *C. jejuni* specific DNA probe. These cultures were grown on Mueller-Hinton agar (Oxoid) plates at 37°C in 7% CO_2, 85% humidity for 1-2 days in a Napco E-series model 5100 CO_2 incubator (National Appliance Co., Tualatin, OR, USA).

A.3. *Escherichia coli* strains

The *Escherichia coli* strains used in this study are listed in Table 2-2. These strains were grown at 37°C either in Miller Luria broth (LB)(Difco) or on Miller LB-plates containing 1.5% agar (Difco). Growth media were appropriately supplemented with antibiotics or other growth factors dependent on the requirements of the procedures used.

E. coli JM83, DH5α and DH5αMCR were used as host strains for pUC20 (Chapter 3). *E. coli* LE392 was the host strain used for cosmids pSa747 and pHC79. *E. coli* KW251 was the host strain for λ-GEM11, while SRB and SRB P2 were used as hosts for λ-DASH.

Solid media used for screening pUC20 plasmids were supplemented with carbenicillin (500 mg/l; Ayerst) or ampicillin (100 mg/l; Ayerst), X-gal (100 mg/l; Gibco/BRL),

Table 2-2. *Escherichia coli* strains used in this research

Strain	Genotype	Source
JM83	*ara, Δ(lac-proAB) rpsL* ϕ80 *lacZM15*	Bethesda Research Laboratories (BRL)
DH10B	F$^-$*mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d *lacZΔM15* Δ*lacX74 deoR recA1 endA1* Δ(*ara,leu*)7697 *araD139 galU galK nupG rpsL*	BRL
ER1793	*fhuA2* Δ(*lacZ*)*r1 supE44 e14- trp-31 his-1 rpsL*104 *xyl-7 mtl-2 metB1* Δ(*mcrC-mrr*) 114::IS10	New England Biolabs
DH5α	F$^-$ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17*(r_K^-,m_K^+) *supE44* λ$^-$*thi-1 gyrA96 relA1*	Gibco/BRL
DH5αMCR	F$^-$*mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoRrecA1 endA1 supE44* λ$^-$*thi-1 gyrA96 relA1*	Gibco/BRL
KW251	F' *supE44 galK2 galT22 metB1 hsdR2 mcrB1 mcrA*$^-$ *agrA81*::Tn10 *recD1014* tetracycline^R	Promega
SRB	*e14*$^-$*(mcrA)* Δ(*mcrCB-hsdSMR-mrr*)*171 sbcC recJ uvrC umuC*::Tn5 *supE44 lac gyrA96 relA1 thi-1 endA1*[F' *proAB lacI*^{q}*ZΔM15*] kanamycin^R	Stratagene
SRB P2	*e14*$^-$*(mcrA)* Δ(*mcrCB-hsdSMR-mrr*)*171 sbcC recJ uvrC umuC*::Tn5 *supE44 lac gyrA96 relA1 thi-1 endA1*[F' *proAB lacI*^{q}*ZΔM15*] spi$^-$ *gam*$^-$ *red*$^-$ kanamycin^R	Stratagene
LE392	F$^-$ *hsdR514* (r_k^- m_k^-) *supE44 supF58 lacYI* or Δ(*lacIZY*)*6 galK2 galT22 metB1 trpR55* λ$^-$	Boehringer Mannheim Company (BMC)

and IPTG (0.5 mM; BMC). For growth and propagation of pUC20 strains and pHC79 strains, solid media was supplemented with carbenicillin (500 mg/l) or ampicillin (100 mg/l). For plasmid and pHC79 DNA isolation, ampicillin (100 mg/l) was added to broth media. Solid and liquid media used for work with cosmid pSa747 were supplemented with 30 mg/l kanamycin (Sigma).

For work with lambda phage, $MgSO_4$ was added to all media, to stabilize lambda phage particles. The $MgSO_4$ was added as a 1.0 M solution through Millex GV 0.22μm micro-filtration filter units (Millipore) using a 10 ml syringe (B-D) to a final concentration of 10 mM. Maltose was also added to top agar and liquid media to a final concentration of 0.5% (w/v), to induce production of the maltose binding protein which acts as the receptor for the phage.

A.4. Vectors

For the construction of genomic libraries of *H. pylori* described in Chapter 3, pUC20 (plasmid, BMC), pSa747 (cosmid, Tait *et al.*, 1983), pHC79 (cosmid, BMC), λ-GEM11 (Promega), and λ-DASH (Stratagene) were used. For subcloning and double stranded sequencing the plasmid pBluescript (Stratagene) was used.

B. Cloning techniques

B.1.Isolation of chromosomal DNA

Genomic DNA was isolated from *Helicobacter pylori* strains UA763, UA764, UA765, and UA802, and *Campylobacter jejuni*

strain UA580 using a modified version of a previously described procedure (Ezaki *et al.*, 1988). Briefly, organisms were grown on appropriate solid media. Cells were collected with a sterile loop and suspended in 1.0 ml of Mueller-Hinton broth in a 1.5 ml microfuge tube. The cells were harvested immediately, without incubation, in an Eppendorf model 5415C microfuge (Brinkman, Brinkman Instruments (Canada) Ltd., Rexdale, ON, Canada) at 16,000 x g for 30 s. The supernatant was removed and the cell pellet was resuspended in 300 µl of isolation buffer (Appendix A). RNase A (1.0 mg/ml; Pharmacia, Pharmacia Biotech Inc., Baie d'Urfe, PQ, Canada) was added to a final concentration of 10µg/ml. Sodium dodecyl sulfate (SDS, 20% w/v solution; ICN, ICN Biochemicals Canada Ltd., Mississauga, ON, Canada) was added to a final concentration of 0.5%. The suspension was incubated at 65°C for 10 min. The suspension was then extracted twice with 500 µl of phenol:chloroform:iso-amyl alcohol (25:24:1) and the aqueous phase was transferred to a new tube. The tube was filled with chilled (-20°C) 95% ethanol (ETOH) and the DNA was spooled out of solution using a drawn-out glass micropipette. The DNA pellet was rinsed in chilled 95% ETOH and dissolved in 300 µl of TE (Appendix A). To remove residual ETOH, 30 µl of 3.0 M sodium acetate (NaOAc) (pH 7.0) and 1.0 ml of chilled 95% ETOH were added. The DNA was recovered by centrifugation in a microfuge for 5 min at 16,000 x g, washed once with chilled 70% ETOH and desiccated. The dry pellet was dissolved in an appropriate amount of TE

to give a final concentration of DNA between 0.1-1.0 μg/μl. The purity of the DNA samples was checked and the concentration of the DNA samples was estimated by electrophoresis on 0.5% agarose gels (Section C.1). The gels were also useful in determining the amount of shearing which may have occurred during the isolation process. If the DNA appeared pure and was not significantly sheared it was used for further experimentation.

B.2. Restriction endonuclease digestion of DNA

Restriction endonuclease digestions were carried out in conditions suggested by the manufacturers of the enzymes. For genomic library construction, chromosomal DNA was either completely or partially digested with the appropriate restriction endonuclease to give ends compatible with the insertion site of the chosen vector.

Complete restriction endonuclease digestion of DNA

For complete restriction endonuclease digestions DNA was digested with 1-5 units of enzyme/μg DNA for 2-16 hr in conditions recommended by the manufacturer of the enzyme. Digestions were stopped with the addition of 1/10th volume 0.5 M EDTA (pH 8.0). In some instances, digestions were immediately used for electrophoretic analysis. For these samples 1/10th volume of loading dye (Appendix A) was added. For the construction of plasmid libraries of *H. pylori* UA763 (Chapter 3), complete digestions of pUC20 DNA or *H. pylori* UA763 DNA were carried out with the enzymes *Bgl*II, *Eco*RI or

*Sst*I. The lambda vectors and pHC79 were digested to completion with *Bam*HI, while the pSa747 was cut with *Bgl*II.

For chromosomal DNA to be used for genomic library construction it was necessary to remove EDTA and any residual enzyme, which would interfere with the ligation reactions. To do this the DNA was precipitated by the addition of 1/10th volume of 3.0 M NaOAc and 3 volumes of chilled 95% ETOH. The DNA was recovered in a microfuge and the pellet was washed twice with 200 µl of chilled 70% ETOH. The DNA was desiccated and the pellet was dissolved in 10 µl of either TE or double-distilled water.

Genomic restriction digestions of the *H. pylori* strains UA763 and UA802 were carried out as described in Chapter 4, using the restriction endonuclease *Bgl*II.

Partial digestion of chromosomal DNA for cloning

Partial digestion of chromosomal DNA was carried out for the construction of genomic cosmid and lambda libraries of *H. pylori* strains UA763 and UA802 (Chapter 3). The genomic DNA was partially digested with either *Sau*3AI or *Bgl*II. Approximately 50 µg of chromosomal DNA were digested in a volume of 125 µl with 0.1-0.5 units of enzyme/µg of DNA under conditions recommended by the manufacturer of the enzyme. At 5, 10, 15, and 30 min 25µl aliquots were removed and added to tubes containing 5µl of 0.5M EDTA (pH 8.0). After 2 hr 5 µl of 0.5 M EDTA (pH 8.0) were added to the final 25 µl of the digestion. The digestions were pooled and 10 µl of 3M NaOAc (pH 7.0) and 300 µl of chilled 95% ETOH were added. The DNA

was precipitated overnight at -20°C. The DNA was recovered by centrifugation in a microfuge and the pellet was washed once with chilled 70% ETOH. The pellet was desiccated and dissolved in 100 µl of TE.

B.3. Sodium chloride gradient size fractionation of DNA

Partially digested DNA was passed through a continuous 13.0 ml 1.0-5.0 M sodium chloride gradient in order to fractionate the DNA fragments by size. The gradients were made from the bottom up using an Easy Cast™ gradient maker (Owl Scientific Plastics Inc.) connected to a model P-1 peristaltic pump (Pharmacia) and mixed with a vibrating mixer (Haakebuchler Instruments Inc.). To form the gradients, 6.9 ml of 1.0 M NaCl in TE were added to the front chamber of the gradient maker and 6.3 ml of 5.0 M NaCl in TE were added to the back chamber. The mixer and the peristaltic pump were switched on and the flow from the back chamber was opened. The centrifuge tube was filled from the bottom up through the peristaltic pump via a plastic tube connected to a siliconized glass FISHERbrand disposable 100 µl micro-pipette (Fisher, Fisher Scientific, Edmonton, AB, Canada). The partial digestions were carefully layered on top of the gradients and the gradients were centrifuged at 38,000 rpm for 4 hr at 22°C in a Beckman L-80 ultracentrifuge (Beckman, Beckman Instruments (Canada) Inc., Mississauga, ON, Canada) using a Beckman model SW40 swinging bucket rotor. To fractionate the gradients, the peristaltic pump was connected

to a fraction collector and 300 µl fractions were collected from the bottom of the tubes into 1.5 ml microfuge tubes. The fractions were de-salted by first precipitating the DNA with 1.0 ml of chilled 95% ETOH. The pellets were then washed several times in 70% ETOH at room temperature until they lost their grainy texture (precipitating salt). The pellets were air dried (5-10 min) and were dissolved in 100 µl of TE. To check the range of fragment sizes of the DNA in each fraction, 10 µl of each fraction were subjected to electrophoresis, with appropriate standards, on a 0.4% agarose gel. For cosmid libraries, DNA fragments of greater than 23 kb were needed. Lambda libraries required DNA fragments in the range of 10-22 kb. The fractions containing DNA fragments in the correct size ranges were re-precipitated by adding 10 µl of 3M NaOAc (pH 7.0) and 300 µl of chilled 95% ETOH. The pellets were dissolved in TE and pooled to give a final volume of 50 µl. This pooled DNA was then used for ligation to either cosmid or phage vectors.

B.4. Preparation of vector DNA for cloning

Preparation of plasmid and cosmid vectors

Plasmid and cosmid vectors were digested to completion with restriction enzymes to give ends compatible with the passenger DNA fragments. In some cases double digestions of vectors were used to give directional clones. If a digestion buffer compatible with both enzymes was available, the digestions were carried out with both enzymes added simultaneously. If no compatible buffer was available, the

DNA was digested with the first enzyme, precipitated, using 1/10th volume NaOAc and 3-4 volumes chilled 95% ETOH, and re-dissolved in buffer suitable for the second digestion. Salt and residual enzymes which could interfere with ligation were removed either by precipitation of the DNA using NaOAc (1/10th volume) and chilled 95% ETOH (3-4 volumes) or by centrifuging the digested vector through an 0.45μm Ultrafree®-probind filter unit (Millipore Corp., Bedford, MA, USA).

Preparation of lambda vectors

The lambda replacement vectors, λ-GEM11 and λ-DASH (see Chapter 3 for description of the vectors), were first digested to completion with restriction enzymes that gave ends compatible with the passenger DNA fragments (*Bam*HI). Right and left arms of the vector were then isolated by electrophoresis through a 0.5% low melting point (LMP) agarose gel, followed by excision and purification of the arms from the LMP gel (Section C.5). The arms were dissolved in TE and pooled in a final volume of 15 μl and used for ligation to passenger DNA. In some cases vector arms which had already been restricted and purified were commercially available.

B.5. Ligation of vector to passenger DNA

For ligation of DNA fragments to either plasmid or cosmid vectors, a passenger:vector molar ratio of approximately 2:1 was used. For ligation of DNA fragments to lambda vectors arms, a passenger:vector arms molar ratio of approximately 1:1 was used. The ligations were carried out at 4°C

overnight, 15°C overnight or room temperature for 2-4 hr in buffer conditions specified by the manufacturer of the enzyme (Gibco/BRL Life Technologies Inc. (BRL), Burlington, ON, Canada). In general, about 1-4 μg of total DNA were used for ligation reactions with 1 unit of enzyme in a final volume of 10 μl.

B.6. Transformation of competent cells

Competent *E. coli* cells obtained commercially or made in our laboratory were used for transformation of plasmid or cosmid vectors. Competent cells were prepared as described (Ausubel *et al.*, 1987). Briefly, the appropriate strain of bacterial culture was inoculated into 100 ml of suitable liquid medium. The culture was grown overnight (16 hr) and harvested in a 250 ml centrifuge bottle at 4°C for 5 min at 5,000 rpm using a Beckman model JA-14 rotor in a Beckman model J 2-21 centrifuge. The cells were then gently resuspended in 10 ml of chilled Ca-buffer (Appendix A) and transferred to a sterile 17 x 100 mm capped polypropylene culture tube. The cells were incubated on ice for 30 min and collected at 4°C for 5 min at 5,000 rpm using a Beckman model JA-20.1 rotor. The cells were gently resuspended in 3.0 ml of a chilled solution of 60 mM $CaCl_2$, 15% (v/v) glycerol, and 10 mM PIPES, pH 7.0. The competent cells were divided into 300 μl aliquots in 1.5 ml microfuge tubes. If the cells were not used immediately, they were flash-frozen in liquid nitrogen and stored at -80°C.

To transform cells, four serial 10-fold dilutions of the ligated DNA were made in 100μl volumes of chilled Ca-buffer to give aliquots containing 1.0 μl,0.1 μl, 0.01 μl, and 0.001 μl of ligated DNA solution. Twenty-five μl of competent cells were added to each tube and the tubes were incubated on ice for 90-120 min. The cells were then heat-shocked at 42°C for 2 min and plated onto suitable solid media, supplemented with antibiotics as required. The plates were incubated at 37°C with the agar layer up (lid down) overnight. The following day, colonies were counted to determine the efficiency of ligation and transformation. Isolated colonies were selected for screening (see Section B.8 below).

B.7. *In vitro* packaging and transfection of lambda and cosmid DNA

Lambda or cosmid DNA was packaged into viable bacteriophage heads before transfection. Several commercially prepared *in vitro* lambda packaging extract kits were available. For this research, extracts from Boehringer Mannheim Canada Ltd. (BMC, Laval, PQ, Canada), United States Biochemical (USB, Cleveland, OH, USA), and Stratagene (La Jolla, CA, USA) were all used following the protocols accompanying the packaging kits. For the packaging reactions 4-6 μg of ligated DNA in a volume of 4 μl were used.

For transfection, 5 ml overnight cultures of appropriate bacterial cells were grown in sterile 17 x 100 mm polypropylene culture tubes in LBM (Appendix A). The cells were harvested at 4°C for 5 min at 5,000 rpm using a Beckman

model JA-20.1. The cells were resuspended in 1 ml of 10 mM $MgSO_4$ and placed on ice. Serial 10-fold dilutions of the packaged DNA suspensions were made in 10 mM $MgSO_4$ and divided into 100 μl aliquots. To each aliquot, 50 μl of cells were added. The cells were incubated for 20 min at 37°C. The cells were then added to 2.5 ml of top agar (culture medium, 10 mM $MgSO_4$, 0.5% maltose, 0.7% agarose held at 48°C) in 13 x 100 mm test tubes, mixed rapidly, and plated on solid media. The top agar was allowed to solidify at room temperature for 10 min and the plates were incubated lids up at 38-40°C overnight. The following day the lambda plaques or cosmid-containing colonies were enumerated to determine the titer of the packaged lambda/cosmid suspensions. Five ml of SM buffer (Appendix A) were overlaid on plates with greater than 200 lambda plaques. These plates were then incubated overnight at 4°C and the following day the suspensions were transferred to sterile 17 x 100 mm capped polypropylene tubes (Fisher), and 1 ml of chloroform was added. These 5 ml suspensions served as lambda library stocks and were re-plated for screening (see Section B.8). Isolated cosmid colonies were screened as described in Section B.8.

B.8. Replica filters for screening

Screening of plasmid and cosmid clones

For the screening of plasmid or cosmid clones, isolated colonies from the primary plates (from Section B.6 or Section B.7 above) were spotted, using a sterile tooth pick, in

arrays onto two plates of solid growth medium. On one plate sterile 82 mm nitrocellulose disks (MSI, Micron Separation Inc., Westboro MA, USA), marked for orientation and in some cases marked with grid lines, were carefully layered onto solid growth media and the colonies were spotted directly onto the nitrocellulose. The second plate did not have a nitrocellulose disk and was used as a master plate. The plates were incubated at the appropriate temperature (usually 37°C) overnight or until the colonies reached approximately 5 mm in diameter. The filter plate was processed by first carefully removing the filter with sterile forceps and placing the filter colony-side up for 5 min on Whatman 3MM chromatography paper which had been soaked in 10% SDS. This step allowed lysis of the cells. The filter was then transferred, colony side up, to Whatman 3MM chromatography paper pre-soaked in a denaturing solution (Appendix A) and then allowed to soak in this solution for 10 min. The filter was then neutralized by floating it on neutralizing solution (Appendix A) for 10 min. The filter was then rinsed in 2 x SSC, air dried and the DNA was fixed to the filter by baking (no vacuum) overnight at 65°C. The filters were screened by hybridization as described below (Section C.4).

Screening of lambda clones

Lambda plates, made as described in Section B.7, were screened by transferring the lambda DNA to nitrocellulose disks. Nitrocellulose disks were marked for identification and orientation, then sterilized by autoclaving. Sterile

forceps were used to layer these disks, with the markings down, onto the plates (ie. in contact with the plate surface). The filters were allowed to soak on the plate for 1 min. During this period the bottoms of the plates were marked to orient the filters. The filters were then carefully removed and floated plaque-side up on denaturing solution for 30-45 s. The filters were then floated on neutralizing solution for 5 min and rinsed in 2 x SSC (Appendix A). The filters were air dried and the DNA was fixed onto the filter by baking (no vacuum) at 65°C overnight. The filters were screened by hybridization as described below (Section C.4). Positive plaques were isolated using sterile toothpicks, and placed in 500 µl of SM buffer with 50 µl of chloroform and stored in this solution indefinitely at 4°C. The positive plaque stock was amplified by re-plating 50 µl of the phage suspension (Section B.7). This amount of phage suspension was normally enough to give confluent plaque formation (ie. the plate was clear). The plate was overlaid with 5 ml of SM (Appendix A) and incubated overnight at 4°C. The overlay was then transferred to sterile 17 x 100 mm capped polypropylene tubes (Fisher) and 1 ml of chloroform was added to each tube. This stock solution could be kept at 4°C indefinitely.

Four plates with greater than 200 plaques of each of the *Bgl*II and *Sau*3AI λ-DASH genomic libraries of *H. pylori* UA802, were screened with PCR probes labelled by nick translation with [^{32}P]-dATP (Chapter 4). Approximately 2000

plaques/library were screened. Positive clones were isolated using sterile toothpicks and stock suspensions were made, as described above. Some of the positive clones were further characterized by restriction digestion analyses and sequence analyses.

B.9. Isolation of plasmid/cosmid DNA

Mini-preparations of plasmid/cosmid DNA

Mini-scale plasmid/cosmid DNA isolations were carried out using methods modified from those of Birnboim and Doly (1979). Five ml broth cultures, supplemented with appropriate antibiotics, were inoculated with single colonies harboring the plasmid/cosmid to be isolated. The cultures were grown in sterile 17 x 100 mm capped polypropylene tubes (Fisher) overnight at 37°C in a tube roller (Model TC-7; NBSC, New Brunswick Scientific Co., Edison, NJ, USA). The cells were harvested in a benchtop centrifuge (Sorvall GLC 2B General Laboratory Centrifuge, DuPont Instruments) for 3 min at 4000 rpm, resuspended in 100 µl of Birnboim & Doly solution 1 (Appendix A) and transferred to 1.5 ml microfuge tubes. Two hundred µl of Birnboim and Doly solution 2 (Appendix A) were added and mixed by inverting the tube. After 1 min, 150 µl of Birnboim and Doly solution 3 (Appendix A) were added and mixed by inversion. The tubes were then centrifuged in a microfuge at 16,000 x g for 15 min. Pelleted material was removed from the tube using sterile toothpicks and the tubes were filled with chilled 95% ETOH. These tubes were then incubated at -20°C for at least 2 hr.

The precipitated DNA was then collected in a microfuge for 2 min and washed twice with chilled 70% ETOH. The final ETOH wash was removed and the DNA pellets were dessicated in a vacuum dessicator (Fisher) for 5 min. The DNA was then dissolved in TE + 1μg/ml RNase A (BMC). In some instances the DNA was further cleaned by phenol/chloroform extraction and precipitation or by passing the DNA solution through a 0.45μm Ultrafree®-probind filter unit (Millipore).

Large scale isolation of plasmid/cosmid DNA

Large scale DNA isolations were done either by use of a DNA isolation kit (Quiagen) or by using methods modified from those of Birnboim and Doly (1979). The modified Birnboim and Doly isolations were carried out as follows. Two hundred ml of culture media, appropriately supplemented with antibiotics, were inoculated with a single colony of bacterial culture harboring the plasmid/cosmid of interest. The cultures were incubated overnight at 37°C in a controlled environment incubator shaker (NBSC). Cells were then harvested in 250 ml centrifuge bottles at 4°C for 5 min at 5,000 rpm using a Beckman model JA-14. The cells were resuspended in 2.5 ml of Birnboim and Doly solution 1 and transferred to a 50 ml polypropylene centrifuge tube. The cell suspension was chilled on ice for 10 min. Five ml of Birnboim and Doly solution 2 were added and mixed by inversion using Parafilm® (American National Can, Greenwich, CT, USA) to cover the tube mouth. After 2 min of incubation on ice, 3.0 ml of Birnboim and Doly solution 3 were added and

the suspension was again mixed by inversion. Incubation was continued on ice for 30 min and the tube was then centrifuged at 12,500 rpm at 4°C for 30 min using a Beckman model JA-20 rotor. The supernatant was transferred to a new 50 ml centrifuge tube and 30 ml of chilled 95% ETOH were added. The tube was mixed by inversion and incubated for at least 2 hr at -20°C. The tube was then centrifuged at 10,000 rpm at 4°C for 10 min using a Beckman model JA-20 rotor. The supernatant was discarded and the pellet was dried in a dessicator for 10 min. The pellet was dissolved in 9.0 ml of TE, and 9.0 g of CsCl were added. The CsCl was dissolved and the samples were then transferred into quick seal centrifuge tubes (Ultra-Clear®, Beckman). Ethidium bromide was added to give a final concentration of approximately 150 μg/ml, the tube was filled to the top with a solution of CsCl (1.0 g/ml in TE), and the tube was sealed. The tubes were then centrifuged for 22 hr at 22°C, at 55,000 rpm using a Beckman Ti70 rotor in a Beckman L-8 ultracentrifuge. The plasmid band was then removed from the centrifuge tube using a 16 gauge needle (B-D, Becton Dickinson and Company, Rutherford, NJ, USA) and a 3.0 ml syringe (B-D), and transferred to a 17 x 100 mm capped polypropylene tube. The ethidium bromide was removed by extraction with iso-amyl alcohol (Sigma) and the DNA was precipitated with 2.5 volumes of chilled 95% ETOH. The pellet was washed several times with chilled 70% ETOH, until the pellet was no longer "grainy". The DNA was then air-dried for 15 min and

redissolved in TE to give a final concentration of approximately 0.25-1.0 μg/μl.

B.10. Isolation of lambda DNA

Preparation of host cells for transfection

Single isolated colonies of the host cells were inoculated into 5 μl of LBM-broth (Appendix A) in sterile 17 x 100 mm capped polypropylene tubes. One tube produced enough cells for 20 mini preparations or one large scale preparation. The cells were grown overnight at 37°C in a tube roller (NBSC). The following morning the cells were harvested in a benchtop centrifuge (Sorvall) for 3 min at 4000 rpm. The cells were then resuspended in 1.0 ml of 10 mM $MgSO_4$ and used for transfection.

Mini-preparations of lambda DNA

In a 1.5 ml microfuge tube, approximately 2.5 x 10^7 plaque forming units (PFU) of phage were combined with 50 μl of host cells in a final volume of 200 μl. The phage were allowed to transfect the host cells at 37°C for 15-20 min and the suspension was then added to 4 ml of LBM-broth in sterile 17 x 100 mm capped polypropylene tubes. The broth culture was grown for 9-12 hr. on a tube roller at 37°C. Three hundred μl of chloroform and 1 ml of 5.0 M NaCl were added and the suspension was incubated for an additional 10 min at 37°C. Cellular debris was removed from the phage suspension by centrifugation at 11,000 rpm at 4°C for 15 min using a Beckman model JA-20.1 rotor. The supernatant was transferred

to a fresh tube and the phage particles were precipitated by the addition of 1.25 ml of 40% w/v polyethylene glycol (PEG) 8000 and incubated on ice for 20 min. The phage particles were collected by centrifugation at 11,000 rpm at 4°C for 15 min. Phage were resuspended in 400 µl of TM (Appendix A) and transferred to 1.5 ml microfuge tubes. Pancreatic DNase and RNase A were added to a final concentration of 1µg/ml and the tubes were incubated for at least 30 min at 37°C. The suspension was extracted twice against one volume of chloroform. The bacteriophage particles were then collected by centrifugation at 60,000 rpm at 22°C for 1 hr in a Beckman TLA-100.2 rotor in a Beckman TL-100 tabletop ultracentrifuge. The supernatant was carefully decanted leaving a "glassy" precipitate at the bottom of the tubes. The pellet was suspended in 300 µl of TM and transferred to a microfuge tube. The centrifuge tube was washed with an additional 100 µl of TM, and wash was pooled with the original phage suspension. Proteinase K (at a final concentration of 50 µg/ml), SDS (at a final concentration of 0.1%) and EDTA (at a final concentration of 20 mM) were added and the suspension was incubated at 65°C for 20 min. The solution was then extracted once with one volume of phenol, once with one volume of phenol:chloroform and once with one volume of chloroform. The DNA was then precipitated by the addition of 1/10th volume of 3 M NaOAc and 3 volumes of chilled 95% ETOH. The tube was incubated for at least one hr at -20°C. The DNA was collected in a microfuge and washed twice with 300 µl of

chilled 70% ETOH. The pellet was allowed to air-dry for about 15 min and was dissolved in an appropriate volume of TE to give a final concentration of approximately 0.2-1.0 μg/μl. The concentration and purity of the DNA were assessed on a 0.6% agarose gel.

Large scale isolation of lambda DNA

The procedure for large scale isolation of lambda DNA was similar to the mini-preparation of lambda DNA but with the following changes. In a 13 x 100 mm test-tube, approximately 6 x 10^8 PFU of phage were used to transfect 1.0 ml of host cells. The transfected cells were then added to 100 ml of LBM in sterile 250 Erlenmeyer flask. The flasks were incubated overnight at 39°C in a controlled environment incubator shaker (NBSC). Two hundred and fifty ml centrifuge bottles were used for removal of cellular debris and centrifugation was carried out at 6,000 rpm at 4°C for 15 min with a Beckman JA-14 rotor. PEG precipitation of the phage particles was carried out as above with scaled-up volumes. The phage were resuspended in 750 μl of TM and transferred to 1.5 ml microfuge tubes. The suspension was treated with pancreatic DNase (1μg/ml) and RNase A (1μg/ml) and extracted twice against one volume of chloroform. The bacteriophage particles were then collected by centrifugation at 60,000 rpm at 22°C for 1 hr in a Beckman TLA-100.2 rotor in a Beckman TL-100 tabletop ultracentrifuge. The phage particles were treated with proteinase K, SDS and EDTA as above and incubated at 65°C for 20 min. Phenol, phenol:chloroform and

chloroform extractions were carried out as described and 1/10th volume of 3 M NaOAc and 3 volumes of chilled 95% ETOH were added. The DNA was spooled out of the tube with a drawn-out glass micropipette and dissolved in 300 μl of TE. The DNA was re-precipitated with 1/10th volume of 3 M NaOAc and 3 volumes of chilled 95% ETOH. The DNA was washed twice with 300 μl of chilled 70% ETOH, air-dried, and dissolved in TE to give a final concentration of approximately 0.2-1.0 μg/μl. The concentration and purity of the DNA were assessed on a 0.6% agarose gel.

C. DNA manipulation techniques

C.1. Agarose gel electrophoresis

Agarose gel electrophoresis was used for a variety of purposes such as restriction digestion analysis of DNA, sizing of intact or fragmented DNA, purity checks of DNA, Southern transfer and hybridization analysis, and analysis of PCR products. The concentration and size of the agarose gels as well as the duration of electrophoresis were determined by the size range of the DNA to be separated. Denaturing agarose gels were used in the analysis of RNA samples (Section F.2). PFGE was used for separation of large DNA fragments for genomic restriction map construction (Section D.3). Low melting point agarose gels were used for extraction of DNA fragments (Section C.5).

Gels were run in 0.5 x Tris-borate (TB) buffer (45 mM Tris base, 45 mM boric acid, 1 mM EDTA, pH 8.3) using horizontal

gel electrophoresis chambers (mini sub-cell, wide mini sub-cell, Bio-Rad, Bio-Rad Laboratories Ltd., Mississauga, ON, Canada; Aquabogue Model 850, Aquabogue Machine Shop, Aquabogue, NY, USA). Bio-Rad model 1450 (Bio-Rad) or Fisher Biotech model FB 105 (Fisher) power supplies were used for electrophoresis. Either a 1 kb ladder (BRL) or a 123 bp ladder (BRL) were used as size standards for linear DNA.

C.2. Southern transfer

DNA was transferred from agarose gels to either supported nitrocellulose (NitroPlus; MSI) or nylon transfer membrane (Hybond; Amersham, Amersham Canada Ltd., Oakville, ON, Canada) by the methods of Southern (Southern, 1975). The DNA was fixed to the nitrocellulose membranes by baking (no vacuum) at 65°C for a minimum of 2 hr. For nylon membranes, the DNA was cross-linked to the membranes in a Stratalinker® UV crosslinker oven (Stratagene).

C.3. Preparation of probes for hybridization

Nick translation was used for labelling probes with radioactive nucleotides. Random priming was used for labelling probes with either non-radioactive digoxygenin (DIG) or with radioactive nucleotides.

Nick translation

The DNA to be labelled (0.3-1.0 μg) was incubated in a 500μl microfuge tube in NT buffer (Appendix A) with 1.0 μl [α^{32}P]-dATP (3000 mCi/mmol, 10 mCi/ml; Dupont Canada Inc., Mississauga, ON, Canada), 2.0 μl of DNase (0.2 μg/ml; BMC) and 1.0 μl of DNA polymerase I (5 units; BMC) in a final

volume of 20 μl for a minimum of 90 min at 15°C in a Lauda K-2/RD circulating water bath (Brinkman). For some labelling experiments [α^{32}P]-dCTP (3000 mCi/mmol, 10 mCi/ml; Dupont) was used instead of [α^{32}P]-dATP. For these experiments dATP replaced dCTP in the 1 X NT buffer. Some probes were labelled using a commercial Nick Translation Kit (BRL) following the protocols specified by the manufacturer of the kit, using the same incubation times and temperatures as above. The labelled DNA was then precipitated by the addition of 1/10th volume of 3.0 M NaOAc (pH 7) and 3 volumes of chilled 95% ETOH (-20°C). The DNA was collected by centrifugation in microfuge for 2 min. The supernatant was carefully removed, placed in an microfuge tube, and discarded to an appropriately shielded radioactive discard receptacle. The pellet was washed once with 80 μl of chilled 70% ETOH and collected in the microfuge. Again the supernatant was discarded to the radioactive discard receptacle. The pellet was allowed to dry for 15 min and redissolved in 100-200 μl of TE. One μl of the labelled DNA was analyzed in a Beckman model LS6800 scintillation counter to determine specific activity.

Random primer labelling

Probe DNA (0.3-1.0 μg) was denatured by heating at 95°C for 10 min in a 500 μl microfuge tube. The denatured DNA was chilled rapidly on wet ice and 2.0 μl of random hexanucleotide mix (0.8 μg/μl; University CORE DNA Services, Calgary, AB, Canada), 2.0 μl 10 X NT buffer (Appendix A),

1.0 μl [α^{32}P]-dATP (3000 mCi/mmol, 10 mCi/ml; Dupont) and 1 μl of Klenow polymerase (1 unit/μl; BRL) were added to the tube and the volume was adjusted to 20 μl with sterile distilled water. For DIG labelling the 10 X NT buffer and [α^{32}P]-dATP were replaced by 2.0 μl dNTP mix (Appendix A). The labelling mixture was then incubated at 37°C for at least 60 min. In most cases the labelling reaction was allowed to incubate overnight. The probe DNA was then precipitated and redissolved in 100μl of TE. For the DIG-labelled probe, 4.0 M LiCl was used in place of the 3.0 M NaOAc. The [α^{32}P]-dATP labelled probes were counted in a scintillation counter.

C.4. Hybridization

Membranes prepared as outlined above (Sections B.8 and C.2) were hybridized with either radioactive or DIG-labelled probes as follows. Membranes were pre-hybridized by immersing the filters in hybridization solution (Appendix A). It was necessary to wet nitrocellulose membranes in 2 X SSC prior to pre-hybridization. Pre-hybridization was carried out either in sealable plastic freezer bags, or shallow sealable plastic containers at 37°C for at least 4 hr for radioactive probes and at least 1 hr for DIG-labelled probes. For hybridization of an average size filter (10 X 15 cm), a volume of 5.0 ml for the sealable bags or 50 ml for the plastic containers was used. Labelled probes were boiled in a water bath for 2-10 min and were then added to the pre-hybridization solution. Generally, 1 x 10^6 CPM of radioactively-labelled probe/filter or $\geq$10 ng/ml of DIG-

labelled probe were added. The hybridization was then allowed to proceed, depending on the desired stringency, at 37-42°C for a minimum of 4 hr.

Processing membranes screened with radioactive probe

Radioactively-probed filters were washed twice after hybridization in 2 X SSC, 0.1% SDS wash solution at 65°C for 30 min. Excessive background signals were removed by washing in 0.2 X SSC, 0.1% SDS at 65°C for an additional 30 min. This process was repeated if the background was still high. The washed membranes were dried and exposed at -82°C to Kodak X-OMAT AR X-ray film (Kodak, Eastman Kodak Co., Rochester, NY, USA) in a shielded X-ray cassette with an intensifying screen (Cronex Lightning plus, Dupont). The exposure times were varied to give the best results with the minimum amount of background. In most cases overnight exposure was adequate.

Processing membranes screened with DIG-labelled probe

Following hybridization, membranes probed with DIG-label were first washed at room temperature in 2 X SSC, 0.1% SDS twice for 5 min. This was followed by two 15 min washes in 0.1 X SSC, 0.1% SDS at 65°C. The membranes were briefly rinsed in wash buffer (Appendix A), then equilibrated for 30 min in blocking buffer (Appendix A). The membranes were then immersed in a solution of anti-DIG-alkaline phosphatase conjugate diluted 1:5000 in blocking buffer for 30 min at room temperature. Approximately 20 ml of diluted conjugate/100 cm^2 surface area of membrane were used. To

remove excess conjugate, the membranes were then washed twice for 15 min at room temperature with wash buffer. The membranes were equilibrated for 5 min at room temperature in PPD-buffer (Appendix A) and immersed in a solution of Lumigen PPD (BMC) diluted 1:100 in PPD-buffer. Following this the membranes were removed and the Lumigen solution was allowed to drain from the membrane. Excess solution was removed by briefly blotting the membranes (DNA side up) on Whatmann 3MM chromatography paper. The membranes were then wrapped in Saran Wrap® or sealed in sealable freezer bags without drying and incubated for 15 min at 37°C. The wrapped membranes were then exposed to Kodak X-OMAT AR® X-ray film in an X-ray cassette at room temperature. Exposure times were varied depending on the strength of the signal and amount of background. Usually exposure times of 15-30 min were sufficient. If membranes were not allowed to dry, they could be re-probed. To re-probe a membrane it was first washed in sterile distilled water, then washed twice at 37°C in a solution of 0.2 M NaOH and 0.1 %SDS. The membrane was rinsed briefly in 2 x SSC and hybridized again as described above.

C.5. Low melting point agarose extraction of DNA fragments

Electrophoresis of DNA was carried out using standard procedures through low melting point agarose gels of appropriate concentrations. The gels were stained with ethidium bromide and the appropriate DNA fragment(s) were excised and excess agarose was trimmed off. The agarose

blocks containing DNA were placed into microfuge tubes with 100-200 µl of TE buffer. The LMP agarose was melted by incubation at 65°C for 10 min. In order to separate the agarose from the DNA sample, the suspensions were extracted twice with one volume of phenol. The aqueous phase was carefully removed and the volume of the DNA solutions was reduced by extracting the aqueous layer with n-butanol (Anachemia Science, Montreal, PQ, Canada). This step also removed any traces of ethidium bromide which might have still been bound to the DNA. The DNA solutions were extracted once against one volume of chloroform and the DNA was precipitated by adding 1/10 volume of 3 M NaOAc and 2-3 volumes of 95% ETOH and incubating at -70°C for 30 min or -20°C overnight. The DNA was collected in a microfuge and the supernatant was discarded. The DNA pellets were washed once with 100 µl of chilled 70% ETOH and centrifuged again. The ETOH wash was removed and the DNA pellets were dessicated. The pellets were redissolved in an appropriate volume of TE.

C.6. Polymerase chain reaction (PCR) methods

PCR was used to generate specific DNA fragments for use as probes (Chapter 4). Conditions were varied depending on the *Taq* polymerase used, the degree of homology of the primer to the template DNA, and the size of the expected PCR product. In most cases, basic conditions, described below, were adequate. Using Promega *Taq* polymerase, the reactions were run in 1 X reaction buffer (provided as a 10 X solution by the manufacturer), with 1.25 mM $MgCl_2$, 0.125 mM dNTP (0.125

mM of dATP,dCTP, dGTP, and dTTP), 20 pmol of each primer, and up to 100 ng of template DNA with 2.5 units of enzyme. Forty cycles of 60s at 94°C, 60s at 45°C and 120s at 72°C were used.

With BRL *Taq* polymerase, the basic conditions were 1 X reaction buffer (provided as a 10 X solution by the manufacturer), with 1.5 mM $MgCl_2$, 0.125 mM dNTP, 20 pmol of each primer, and up to 100 ng of template DNA with 2.5 units of enzyme. Forty cycles of 40s at 94°C, 40s at 45°C and 90s at 72°C were used.

PCR was used to generate probes for the P-type ATPase as well as the 16S and 23S genes of *H. pylori* UA802 (chapter 4). Promega and BRL *Taq* polymerase were used for the amplification in conditions described above. The primers used are described in Chapter 4.

C.7. Partial digestion mapping of lambda clones

Rapid construction of crude restriction maps of lambda clones was accomplished using partial digestion mapping. Briefly, lambda clones were first digested to completion with the enzyme to be mapped, to determine the sizes of the digested fragments of the cloned DNA. The clone was then partially digested with the enzyme, fractionated on 0.5% agarose gels (Section C.1), and the DNA was transferred to nylon or nitrocellulose membranes by Southern transfer (Section C.2). These blots were hybridized with purified right arms of the lambda vector (Section C.5) labelled with either radioactive or non-radioactive label (Section C.3).

The resulting ladder of partially digested fragments which hybridized with the right arm of the vector was used to determine the order of the fragments in the lambda clones, and crude restriction maps were constructed.

D. Pulsed Field Gel Electrophoresis

D.1. Preparation of blocks for PFGE

Pulsed field gel electrophoresis is a method by which large DNA fragments, up to several hundred kb, can be resolved on an agarose gel. To prevent shearing of the large DNA fragments it was necessary to first immobilize bacterial cells in agarose blocks. Bacterial cells from a single plate of pure culture were suspended in 1.0 ml of suspension buffer (Appendix A) in a 1.5 ml microfuge tube. The cells were harvested for 20 s in a microfuge and resuspended in 0.75 ml of suspension buffer. The cells were then added to a 1.5% agarose solution (1.5% agarose in 0.25 M EDTA, 10 mM Tris·HCl pH 7.5, kept at 58°C) and dispensed into a PFGE block mold. The blocks were allowed to solidify at room temperature for 10 min. The blocks were then transferred from the mold into 5 ml of lysozyme buffer (Appendix A) in a capped 17 x 100 mm polypropylene tube and incubated at 37°C for 2 hr. The blocks were then transferred to 5 ml of deproteination solution (Appendix A) in a capped 17 x 100 mm polypropylene tube and incubated at 58°C for 1-2 days. Residual proteinase activity was removed by three 15 min washes in PMSF-buffer (Appendix A) at room temperature. The blocks were then

washed 10-15 times in 5 ml of PFGE block-washing buffer for 20 min at room temperature.

D.2. Restriction digestion of embedded DNA

The washed blocks containing the embedded DNA were sliced into 2 x 2 mm slices using a clean razor blade. Slices were pre-soaked in 400 µl of the appropriate restriction enzyme buffer at room temperature for 15 min. The enzyme buffer was removed, 400 µl of fresh buffer was added, and the slices were incubated for an additional 15 min. This process was repeated 5 times. After the final soak, the enzyme buffer was removed and 400 µl of enzyme buffer with 2 unit/slice of restriction enzyme were added. The digestions were then incubated at the appropriate digestion temperature for 6-16 hr.

D.3. PFGE gels

Gels were run using either the LKB 2015 Pulsaphor plus PFGE system or the Beckman Gene Line II Transverse Alternating Field Electrophoresis (TAFE) system. The concentration of agarose used for PFGE was either 0.75% or 1.0% in PFGE running buffer (Appendix A). For the TAFE, 1.0% agarose gels were used in TAFE running buffer (Appendix A). The gels were cast in molds provided by the manufacturers. The agarose blocks containing the embedded DNA were placed into wells formed in the gels and sealed into place using 1.5% low melting point agarose. PFGE gels were run in PFGE running buffer with constant voltage set at 190v. The gels were run at 8°C and pulse-times were varied from 5 s to 60 s. Pulse

times were varied directly with the size of the fragments to be separated. The duration of the pulse times varied depending on the range of fragment sizes in a given preparation. Once electrophoresis was complete the gels were processed as described in earlier sections either for Southern blots (Section C.2) or for LMP-agarose fragment extraction (Section C.5). For Southern blots of PFGE gels, the gels were first immersed in 0.25 M HCl for 15 min prior to the normal denaturation/neutralization steps.

E. DNA sequencing

E.1. Double-stranded dideoxy-sequencing

Double-stranded dideoxy-sequencing reactions were carried out using [α-^{35}S]-dATP as described (Wang, 1988) or following the protocols included in the Sequenase® dideoxy-sequencing kit (USB). Approximately 2-6 μg of DNA were used for each sequencing reaction. Using these methods the DNA sequences of the cloned DNA from the two *Campylobacter* probes, pDT1719 and pDT1720, were determined (Chapter 6).

E.2. Cycle sequencing with ^{33}P

Cycle sequencing was carried out with [γ-^{33}P]-ATP according to protocols included with the dsDNA Cycle Sequencing System (BRL). The reactions were carried out in a Perkin Elmer model 480 DNA thermocycler (Perkin Elmer) for forty cycles of 30s at 94°C, 30s at 50°C, and 40s at 72°C.

E.3. Sequence analysis

Sequencing gels were run in TBE-buffer (Appendix A) using either the BRL model S1 sequencing gel electrophoresis apparatus or the Bio-Rad Sequi-gen II sequencing cell system. The gels were composed of 5% acrylamide (19:1 acrylamide:bis-acrylamide; Bio-Rad) and 8 M urea in TBE-buffer. To catalyze the solidification of the gels, ammonium persulfate to a final concentration of 0.5 mg/ml and N,N,N',N'-Tetramethylethylenediamine (TEMED) to a final concentration of 0.5 µl/ml were added immediately prior to casting. Gels were run with constant power using either an EC703 power supply (E-C Apparatus Corp., St. Petersburg, FL, USA) or an LKB 2297 macrodrive S power supply (LKB). Gels were pre-run before samples were loaded to allow the gels to reach an optimal temperature for separation, usually between 45-55°C. Samples were denatured at 95°C for 2-5 min immediately prior to loading.

Once the samples were run for the desired time, the gels were immobilized on sheets of Whatman 3MM chromatography filter paper and dried at 80°C for 1 hr on a model SE1160 slab gel dryer (Hoefer Scientific, San Francisco, CA, USA). The dried gels were placed in X-ray film cassettes and exposed to Kodak XAR5® (Kodak) or BioMax® (Kodak) X-ray film, usually overnight. In some instances longer exposures were necessary.

The sequence data was read from the developed film and entered into either the Microgenie computer analysis system (Beckman) or the Wisconsin Sequence Analysis Package

(Genetics Computer Group (GCG), Madison, WI, USA) for computer analysis. Computer analysis of the rRNA gene sequences in Chapter 4, was carried out using the Wisconsin Sequence Analysis Package. Phylogentic comparisons were carried out using the pileup program of the Wisconsin Sequence Analysis Package. This program will align multiple sequences by the progressive alignment method of Feng and Doolittle (1987). The pileup program is also able to plot the comparisons in the form of dendograms. In Chapter 6, the DNA sequence of the two *Campylobacter* probes, pDT1719 and pDT1720, was entered and assembled using Microgenie. Analyses of the sequences and homology searches were carried out with the Basic local alignment search tool (BLAST; Altshul *et al.*, 1990).

E.4. Construction of deletion mutants for sequencing

Deletion mutations of some clones of the 23S/5S genes were carried out to obtain sequential deletions of the 23S/5S genes for sequencing. The deletions were generated using the Promega Erase-a-base system following protocols recommended by the manufacturer with the following changes. Instead of 25 30-second time points, 6 30-second time points were used for deletions. Reagent volumes were adjusted accordingly. Deletion reactions were carried out at 34°C with 400 units of enzyme/5μg of DNA. At this temperature 250-350 bases/min were deleted. A reaction volume of 30 μl was used and 5 μl aliquots were taken for each time point and added to 15μl of

S1 nuclease mix. Eighty µl of ligase mix were then added to each sample.

F. RNA techniques

F.1. Isolation of total RNA

Total RNA was isolated using a modification of the methods of Konkel *et al.* (1994). All solutions used in this section were made up in deionized water treated with diethyl pyrocarbonate (DEPC; Aldrich Chemical Co., Milwaukee, WI, USA; refer to Appendix A under DEPC-water). Briefly, *H. pylori* cultures were spread on solid media and grown as described above. A sterile loop was used to resuspend the cells in 500 µl of protoplast buffer (Appendix A). The suspended cells were divided into two microfuge tubes and 250 µl of 5% SDS were added. The cell suspensions were incubated for 15 min at 70°C. The cell suspensions were then cooled to room temperature and extracted twice with 300 µl phenol. The aqueous layer was then extracted once with 300 µl phenol:chloroform (1:1) and twice with 300 µl chloroform. Following this, 250 µl of 5.0 M NaCl were added and the suspensions were incubated on ice for 10 min. The precipitate was removed by centrifugation for 2 min in a microfuge and the supernatant was divided into two new microfuge tubes. One ml of chilled 95% ETOH was added and the DNA was spooled out, rinsed in 95% ETOH and dissolved in 300 µl of TE. The remaining supernatant was incubated at -20°C for at least 2 hr. The RNA was collected in a

microfuge and allowed to air dry. The RNA pellets were then dissolved in a final volume of 200 µl of DEPC-water. This solution was then treated with RNase-free DNase I (BMC) and the enzyme was removed by centrifugation of the samples through a Millipore ultrafree®-probind 0.45 µm filter unit. The samples were de-salted by centrifugation through Chroma-spin 100 columns (Clontech, Palo Alto, CA, USA) equilibrated with DEPC-water.

F.2. Denaturing RNA agarose gels

The purity of the RNA samples was checked by electrophoresis in 1.2% denaturing agarose gels (Appendix A) in 1 X MOPS buffer (Appendix A). RNA samples were prepared for electrophoresis by adding 10 µl of denaturing mix (Appendix A) and 2 µl of loading dye to 5 µl of RNA sample and incubating at 65°C for 10 min. After electrophoresis the gel was stained in ethidium bromide for 20 min and destained in dH_2O for at least 1 hr. The RNA could then be transferred to nylon membranes, following normal Southern transfer procedures, without denaturation or neutralization. The RNA was cross-linked to the nylon membrane in a Stratalinker® UV crosslinker oven (Stratagene).

F.3. Primer extension

Primer extension was carried out as described by Newnham and Taylor (1994), using superscript II reverse transcriptase (BRL) with the following changes. Instead of incorporating [α-^{35}S]-dATP, [γ-^{33}P] end-labelled primers were used. End-

labelling was carried out as described in the Cycle Sequencing System (BRL) protocol.

G. SDS-PAGE and Western blots

Crude preparations of the MOMP of *C. jejuni* UA580 were subjected to SDS-PAGE analysis and proteins were blotted to PVDF membranes. SDS-PAGE was carried out as previously described (Laemmli, 1970). Samples were electrophoresed through a 5% stacking gel (Appendix A) and a 10% separating gel (Appendix A) in a Bio-Rad Mini-PROTEAN II cell in SDS-PAGE running buffer (Appendix A). Prior to loading the protein samples, 1/10th volume of SDS-PAGE running dye (Appendix A) was added to the samples and the samples were boiled for 1-2 min. Proteins were run with constant current of no greater than 40 mA. When sufficient separation was obtained, the proteins were then electro-blotted to PVDF membranes (Matsudaira, 1987). Electro-blotting was carried out in a Bio-Rad apparatus in CAPS buffer (Appendix A) with constant current (300 mA) for 30 min. The membrane was then immersed in fresh staining solution (Appendix A) for 5 min. The membrane was then destained with several changes of destaining solution (Appendix A), rinsed with dH_2O several times and allowed to dry in air. The stained membrane could then be stored in a sealable plastic freezer bag indefinitely at -20°C. The stained PVDF membranes were sent to the Alberta Peptide Institute (API, Department of Biochemistry, University of Alberta, Edmonton, Canada) and the 45 kDa MOMP

protein was excised from the membrane for N-terminal amino acid sequence analysis.

Chapter 3

Construction of genomic libraries of *H. pylori*.

A. Introduction

An objective of this research was to construct genomic libraries of *H. pylori* DNA which would provide a source of cloned *H. pylori* DNA fragments from which specific genes or regions of DNA could be isolated. Such libraries will become increasingly useful as the mechanisms of pathogenesis of *H. pylori* and the factors involved are delineated. As discussed in Chapter 1, a high degree of genomic variability has been observed between different strains of *H. pylori*; this variability may be related to the different clinical symptoms of *H. pylori* infection. Isolation and comparisons of virulence factors from various strains may provide us with information about the mechanisms of the different clinical presentations of *H. pylori* infection (ie. gastritis, ulceration). The availability of genomic libraries for different strains of *H. pylori* would facilitate these studies. In addition the genomic libraries can be used to clone other genes or regions of interest, such as the rRNA genes, to further our understanding of this organism.

A major consideration for the construction of genomic libraries was the selection of a vector system. A wide variety of plasmid, cosmid and lambda vectors are available, each having its' own advantages and disadvantages. The selected vector system should produce genomic libraries which

would provide a convenient source of DNA for isolation and characterization of genetic elements.

A.1. Plasmid vectors

There are many convenient plasmid vectors, with a variety of selection and screening markers, that are commercially available. These vectors are useful for efficient cloning of fragments from 0-10 kb in size (Maniatis *et al.*, 1982). The number of clones needed to statistically represent 99% of a genome (***N***) can be calculated using the formula

$$N = \frac{[\ln(1 - \text{probability } (99\%))]}{[\ln(1-(\text{ave.insert size} \div \text{total genome size}))]}$$

(Maniatis *et al.*, 1982). One assumption made in this formula is that all parts of the genome are equally represented in the array of fragments to be cloned. In practice this is not always true. Toxicity of cloned elements and the restriction modification system of the host strain may make some regions difficult to clone. In addition, larger fragments may not be cloned as efficiently and could be poorly represented. To circumvent these problems, construction of plasmid libraries using a number of different restriction enzymes would be necessary.

The size of the *H. pylori* strain UA802 genome has been estimated to be approximately 1.7 Mb, as determined by Taylor *et al.* (1992a) using PFGE. By using the above formula, with the average insert size as 5 kb and the genome as 1.7 Mb, a minimum of 1563 clones would be needed to represent 99% of

the genome. This figure (1563) represents the optimal number of clones needed assuming all regions of the chromosome were represented equally. As mentioned, this is not always the case, and more than one library would likely need to be constructed, using a variety of restriction enzymes. An increase in the number of clones screened, beyond what is needed to statistically represent the genome, would decrease the chances of missing some of the regions that are difficult to clone.

Because a minimum of 1563 clones is needed to represent the genome, and because of the logistical problems of long-term storage of such a large number of clones, plasmid libraries are generally used for cloning specific genes. Since the objectives of this research included the isolation of specific genes such as the P-type ATPase and the rRNA genes, the construction of plasmid libraries was attempted.

A.2. Cosmid vectors

Cosmid vectors are very convenient for the rapid construction of genomic libraries, as large fragments (up to 50 kb) of DNA can be cloned. Cosmid vectors combine the convenience of plasmid vectors with the cloning capacity of lambda vectors by incorporating the lambda *cos* (cohesive end) site into plasmid vectors. The *cos* site is the signal recognized by the lambda packaging machinery which packages DNA of appropriate length into phage heads. The maximum and minimum sizes of DNA which can be packaged to produce viable phage particles are approximately 53 kb and 37 kb

respectively (Chauthaiwale *et al.*, 1992). The large size of the fragments which can be cloned in cosmid vectors means that far fewer clones are needed to statistically represent a given genome. However, the larger size of fragments also increases the chance of cloning toxic regions. Thus, as with the plasmid vectors, the chromosome may not be equally represented.

Since far fewer clones are needed to represent the genome using cosmid libraries, these libraries can easily be cryogenically stored and re-screened at later dates with probes as they become available. In fact, in a recent review, Fonstein and Haselkorn (1995) recommended the use of cosmid vectors for construction of genomic "encyclopedias" of organisms which could be used as convenient sources of DNA for isolation and characterization of genes.

A.3. Lambda vectors

Many lambda cloning systems are now available commercially. These vectors are classified as either insertion or replacement vectors. With insertion vectors, the passenger DNA is inserted into a non-essential region of the vector. Some vectors include the *lacZ*/blue-white selection cassette for added convenience in screening. Since the maximum size of DNA which can be packaged to produce viable phage particles is approximately 53 kb (Chauthaiwale *et al.*, 1992), insertion vectors are limited to cloning DNA fragments of 0-10 kb (Maniatis *et al.*, 1982). The minimum number of clones needed to represent 99% of a genome is the same as

with plasmid vectors (1563). However, the screening, storage and handling of lambda libraries is easier than plasmid or cosmid libraries.

The size of inserts which can be cloned in replacement vectors is significantly greater than can be cloned in insertion vectors. This is because the passenger fragments replace a non-essential region of the lambda DNA which is approximately 14 kb in size. Thus, inserts of 9-23 kb can be cloned into most replacement vectors. This makes them quite convenient for the construction of genomic libraries, as the number of plaques to be screened is reduced when compared with the insertion vectors (487 compared to 1563). In addition, lambda libraries can be stored as phage suspensions for several years, and can be screened repeatedly with probes as they become available.

B. Materials and Methods

Please refer to Chapter 2 for detailed methodology used in library construction.

C. Results

C.1. Construction of plasmid libraries of *H. pylori* UA763

Initial work on the production of genomic libraries of *H. pylori* DNA was conducted using the plasmid vector pUC20. Chromosomal DNA from *H. pylori* UA763 was digested with the restriction enzymes *Bgl*II or *Sst*I. This DNA was cloned into

pUC20 and transformed into a variety of *E. coli* host strains. Table 3-1 shows the results of transformation of *E. coli* host strains DH5α or DH5α MCR with pUC20 DNA ligated with *H. pylori* UA763 DNA. Transformations were repeated several times and the best results obtained are presented. The *E. coli* DH5α and DH5α MCR competent cells were obtained commercially from BRL. The expected transformation efficiencies of these competent cells were 10^6-10^8 colonies/μg of plasmid DNA (BRL), while the competent cells made in the laboratory were expected to give efficiencies of 10^6-10^7 colonies/μg of plasmid DNA (Maniatis *et al.*, 1982). Construction of plasmid libraries of *H. pylori* UA763 was also attempted using *Eco*RI, however, for reasons which could not be determined, no recombinants were observed.

With the *Bgl*II and *Sst*I pUC libraries transformed into the DH5α host, the number of transformants obtained was extremely low (120-500 and 160-400 colonies/μg transformed DNA respectively; Table 3-1). These white colonies were screened with radioactively-labelled *H. pylori* UA763 DNA, and only a few of the clones hybridized (5/42 *Bgl*II colonies using the *E. coli* DH5α strain; 7/65 *Sst*I colonies using the *E. coli* DH5α strain; 57/552 *Bgl*II colonies using the *E. coli* DH5αMCR strain; 52/723 *Sst*I colonies using the *E. coli* DH5αMCR strain). Mini-preparations of plasmid DNA showed that the non-hybridizing white colonies harbored plasmids equal in size to or slightly smaller than wild-type pUC20

Table 3-1. Transformation efficiency of plasmid DNA in *E. coli* DH5α and DH5αMCR strains

Insert DNA[a]	Host strain	No. of colonies[b]	Amount of DNA used for transformation	Transformants per μg DNA
None- *Eco*RI + control	DH5α	342[c]	0.1 ng	3.42×10^6
None- *Eco*RI - control	DH5α	8.0[c]	100.0 ng	80
None- *Eco*RI + control	DH5α MCR	271.0[c]	0.1 ng	2.71×10^6
None- *Eco*RI - control	DH5α MCR	13.0[c]	100.0 ng	130
H. pylori UA763 *Eco*RI	DH5α	0	0.1 ng	0
H. pylori UA763 *Eco*RI	DH5α	0	1.0 ng	0
H. pylori UA763 *Eco*RI	DH5α	0	10.0 ng	0
H. pylori UA763 *Bgl*II	DH5α	0.3 (0)	1.0 ng	3.0×10^2
H. pylori UA763 *Bgl*II	DH5α	5.0 (1)	10.0 ng	5.0×10^2
H. pylori UA763 *Bgl*II	DH5α	12.0 (4)	100.0 ng	1.2×10^2
H. pylori UA763 *Sst*I	DH5α	1.6 (0)	1.0 ng	1.6×10^3
H. pylori UA763 *Sst*I	DH5α	4.0 (1)	10.0 ng	4.0×10^2
H. pylori UA763 *Sst*I	DH5α	18.0 (6)	100.0 ng	1.8×10^2
H. pylori UA763 *Bgl*II	DH5α MCR	56.0(19)	1.0 ng	5.6×10^4
H. pylori UA763 *Bgl*II	DH5α MCR	128.0(38)	10.0 ng	1.28×10^4
H. pylori UA763 *Bgl*II	DH5α MCR	too many to count	100.0 ng	---
H. pylori UA763 *Sst*I	DH5α MCR	68.0(19)	1.0 ng	6.8×10^4
H. pylori UA763 *Sst*I	DH5α MCR	173.0(33)	10.0 ng	1.73×10^4
H. pylori UA763 *Sst*I	DH5α MCR	too many to count	100.0 ng	---

a- pUC20 DNA was cut with appropriate endonuclease and ligated to insert DNA. No ligase was added to -controls

b- average of three plates. Numbers in brackets represent the number of colonies which hybridized with labelled *H. pylori* DNA

c- number of blue colonies on a single plate were counted for the negative controls

with no insert. This indicated that either small deletions were occurring in the multiple cloning site of pUC20 or that the *H. pylori* DNA was not stable and was being deleted.

Use of the commercially prepared competent DH5αMCR strain allowed for a greater transformation efficiency than the other *E. coli* host strains. This strain has been highly altered genetically to reduce the effects of the restriction modification barrier. Even so, the number of clones which carried *H. pylori* DNA was far below the expected number of transformants/μg of DNA. The manufacturer of these competent cells states that the optimal transformation of wild-type, uncut pUC20 would yield 1 x 10^8 colonies/μg of transformed DNA (BRL catalog 1993-94). The values obtained for the re-ligation of digested vectors (positive control) was 2.7 X 10^6 colonies/μg. This was within the acceptable range of 10^6-10^8 colonies/μg of plasmid DNA (BRL catalog 1993-94). The number of recombinants observed for the genomic libraries were, in contrast, 4-6 orders of magnitude lower than optimal.

C.2. Construction of cosmid libraries of *H. pylori* DNA

Since no efficient plasmid libraries could be obtained, construction of cosmid libraries was attempted. For this research the cosmids pSa747 (Tait *et al.*, 1983) and pHC79 (BMC) were used to construct genomic cosmid libraries of *H. pylori*. Table 3-2 shows the highest transfection efficiency obtained using these cosmids. Variation of the vector:passenger ratios, variation of the amounts of DNA

Table 3-2. Efficiency of cosmid libraries in *E. coli* LE392

Insert DNA[a]	Vector	No. of colonies hybridizing to *H. pylori* DNA[b]	Amount of packaged DNA used/plate for transfection	CFU[c]/μg DNA
H. pylori UA763 *Bql*II	pSa747	4 (37)	3 μl (30 ng)	4.11×10^2
H. pylori UA763 *Sau*3AI	pSa747	8 (72)	3 μl (30 ng)	4.66×10^2
H. pylori UA802 *Bql*II	pSa747	7 (56)	3 μl (30 ng)	6.22×10^2
H. pylori UA802 *Sau*3AI	pSa747	11 (93)	3 μl (30 ng)	1.03×10^3
H. pylori UA763 *Bql*II	pHC79	18 (132)	3 μl (30 ng)	1.47×10^3
H. pylori UA763 *Sau*3AI	pHC79	2 (147)	3 μl (30 ng)	1.63×10^3
H. pylori UA802 *Bql*II	pHC79	11 (133)	3 μl (30 ng)	1.48×10^3
H. pylori UA802 *Sau*3AI	pHC79	12 (184)	3 μl (30 ng)	2.04×10^3

a- partial digests of genomic *H. pylori* DNA were ligated to vector cut with *Bam*HI
b- total of 3 plates. Numbers in brackets represent total number of colonies observed
c- colony forming units

packaged, and purification of ligated DNA did not improve these results. With wild type lambda packaged *in vitro*, the optimal expected plaque forming units (PFU) would be 1 x 10^9 plaques/μg of packaged DNA (Stratagene catalog 1992, p.42). The transfection frequencies were far below what were expected for optimal transfection (133-267 CFU/μg packaged DNA for pSa747 clones; 67-600 CFU/μg packaged DNA for pHC79 clones; Table 3-2).

C.3. Construction of lambda genomic libraries of *H. pylori* DNA

The construction of lambda libraries of *H. pylori* DNA was started when problems in cosmid library construction became apparent. Initially the lambda vector λ-GEM11 (Promega) was used to construct libraries of *H. pylori* UA763 and UA802 genomic DNA, however, the efficiency of these libraries seemed low. The vector λ-DASH II (Stratagene) was suggested as an alternative lambda vector which had been used successfully in other labs (Dr. J. Elliot, personal communication). This vector was used to construct genomic libraries of *H. pylori* UA802. *H. pylori* UA802 was selected because a genomic restriction map of this strain had been constructed in our laboratory (Taylor *et al.*, 1992a).

Transfections of the λ-GEM11 libraries into the *E. coli* strain KW251 gave unsatisfactory results (258-410 PFU/μg packaged DNA for λ-GEM11; Table 3-3). In addition to the low number of plaques observed, the size of the plaques was smaller than those observed on control plates (0.5 mm

Table 3.3. Transfection efficiency of lambda libraries

Insert DNA[a]	*E. coli* host strain	Vector	No.of plaques[b]	PFU[c]/μg packaged DNA
H. pylori UA763 *Bgl*II	KW251	λ-GEM11	138	276
H. pylori UA763 *Sau*3AI	KW251	λ-GEM11	205	410
H. pylori UA802 *Bgl*II	KW251	λ-GEM11	156	312
H. pylori UA802 *Sau*3AI	KW251	λ-GEM11	129	258
H. pylori UA802 *Bgl*II	SRB/ SRB-P2	λ-DASH II	378[d]	1.23×10^4
H. pylori UA802 *Sau*3AI	SRB/ SRB-P2	λ-DASH II	366[e]	1.22×10^4

a- partial digests of *H. pylori* genomic DNA were ligated with *Bam*HI lambda vector arms
b- average of 3 plates. For each plate 50 μl (~500 ng packaged DNA) of undiluted phage suspension were used for transfection
c- plaque forming units
d- average of 10 plates. For each plate 3 μl (~30 ng packaged DNA) of undiluted phage suspension were used for transfection
e- average of 4 plates. For each plate 3 μl (~30 ng packaged DNA) of undiluted phage suspension were used for transfection

compared to 1-2 mm). Plaque-lifts were hybridized with radioactively-labelled *H. pylori* DNA. Because of the small size and low number of the plaques, the results were difficult to interpret. It did appear that all of the plaques hybridized with the radioactively-labelled *H. pylori* DNA. The figures presented in Table 3-3 for the λ-GEM11 libraries represent the best values obtained with this lambda system. A number of parameters were varied during these attempts to make a library including variation of vector:passenger ratios, adjustment of total amounts of DNA to be packaged, and purification of ligated DNA through DNA purification columns (Chroma-spin TE 1000), all with similar results.

Results for the λ-DASH II libraries of *H. pylori* UA802 transfected into *E. coli* SRB were considerably better than those observed for the λ-GEM11 libraries (1.22-1.23 x 10^4 PFU/μg packaged DNA; Table 3-3). The size of the plaques observed (1-3 mm) was also increased. In addition, hybridization of plaque-lifts of these plates clearly showed that all of the plaques contained *H. pylori* DNA. The numbers obtained were still far below those expected for construction of lambda libraries. As mentioned above, an optimal efficiency of 1 x 10^9 plaques/μg of packaged DNA was expected. Even though the values obtained were low, there were sufficient PFU to statistically represent several genomes. A minimum of 487 clones was calculated to represent

99% of the genome, and approximately 2000 plaques were obtained for each library.

D. Discussion

D.1. Plasmid libraries

Since downstream characterization of genes cloned in plasmid vectors is much more convenient than in cosmid or lambda vectors, construction of plasmid libraries of *H. pylori* UA763 was attempted. However, no efficient libraries using pUC20 as a vector were produced from this research. The results showed that less than 10% of the colonies contained *H. pylori* DNA. The highest transformation rates for the plasmid libraries of *H. pylori* DNA were obtained using the *E. coli* host strain DH5αMCR. This strain carried mutations of the restriction modification system (Table 3-4). This is the system by which foreign DNA is recognized by the host and destroyed. In *E. coli* several genes are involved in this complex system (Bickle and Krüger, 1993). Thus, it appeared that the restriction modification barrier was one of the factors responsible for difficulties observed in construction of plasmid libraries. Even with the DH5αMCR *E. coli* host strain, however, the efficiency of library construction was still low, suggesting that other factors could also be involved.

These results correlated with those obtained in a recent study of the choice of host strains for the construction of *H. pylori* genomic lambda libraries (Phadnis *et al.*, 1993). The *mcr*C-*hsd*RMS-*mrr* loci seemed to be of particular

Table 3-4. Restriction phenotypes of some *E. coli* strains

Strain	Restriction Phenotype	Source
JM83	*mcr*A$^+$ *mcr*BC$^+$ *hsd*R$^+$ *hsd*M$^+$ *hsd*S$^+$ *mrr*$^+$	BRL
DH5α	*mcr*A$^+$ *mcr*BC$^+$ *hsd*R$^-$ *hsd*M$^+$ *hsd*S$^+$ *mrr*$^+$	BRL
DH5αMCR	*mcr*A$^-$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^-$ *hsd*S$^-$ *mrr*$^-$	BRL
LE392	*mcr*A$^-$ *mcr*BC$^+$ *hsd*R$^-$ *hsd*M$^+$ *hsd*S$^+$ *mrr*$^+$	BMC
KW251	*mcr*A$^-$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^+$ *hsd*S$^+$ *mrr*$^+$	Promega
SRB	*mcr*A$^-$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^-$ *hsd*S$^-$ *mrr*$^-$	Stratagene
SRB P2	*mcr*A$^-$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^-$ *hsd*S$^-$ *mrr*$^-$	Stratagene
HB101	*mcr*A$^+$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^-$ *hsd*S$^-$ *mrr*$^-$	BRL
ER1793	*mcr*A$^+$ *mcr*BC$^-$ *hsd*R$^-$ *hsd*M$^-$ *hsd*S$^-$ *mrr*$^-$	NEB

*mcr*A - encodes restriction of DNA with cytosine methylated at the sequence 5'-C^{me}CGG-3'(Raleigh and Wilson, 1986)

*mcr*BC- encodes restriction of DNA with cytosine methylated at the sequence 5'-G^{me}C-3' (Raleigh and Wilson, 1986)

*hsd*R - encodes *Eco*K and *Eco*B restriction of unmodified DNA (Gough and Murray, 1983)

*hsd*M - encodes the host-specific methylation of DNA (Gough and Murray, 1983)

*hsd*S - encodes the DNA sequence-specificity determinant for *hsd*M and *hsd*R (Gough and Murray, 1983)

mrr - encodes restriction of DNA with adenosine methylated at the sequences 5'-C^{me}AG-3' or 5'-G^{me}AC-3' (Waite-Rees *et al.*, 1991)

importance, as lambda libraries constructed in strains carrying mutations in these alleles had the highest efficiency. These libraries produced approximately 2000 plaques from 5 µl of undiluted phage suspension. Although it was not stated in the study, standard packaging reactions require up to 5 µg of DNA in total and yield approximately 500 µl of suspension. Most commercial *in vitro* packaging systems will give 1 x 10^7-1 x 10^9 plaques/µg of packaged DNA. Thus, the observed efficiency of the libraries constructed by Phadnis *et al.*, using the *E. coli* host strain ER1793, could be estimated to be 8 x 10^4 - 4 x 10^5 plaques/µg packaged DNA, which was still far below the expected efficiency for such a library. This suggested, as did the results of this research, that other factors must also have contributed to decreased library efficiency.

The results of the cloning of *H. pylori* DNA presented here were similar to those obtained with *Streptococcus pneumoniae* DNA. In the construction of genomic libraries of *S. pneumoniae*, it was found that cloned *S. pneumoniae* DNA was highly unstable in *E. coli* when using high copy number vectors such as pBR325 (Chen and Morrison, 1987). Morrison and Jaurin (1990) first suggested that the low G + C content of the *S. pneumoniae* DNA (39 mol.%; Marmur and Doty, 1969) resulted in a high incidence of *E. coli* promoter-like sequences, which had strong enough promoter activity to interfere with the stability of the cloned DNA. This hypothesis was later contradicted in a study showing that

promoters of sufficient strength to cause plasmid instability were probably rare and that even sequences resembling relatively strong *E. coli* promoters did not seem to affect clone stability (Dillard and Yothers, 1991). Although sequence-related effects could not be ruled out in the later study, Dillard and Yothers concluded that problems associated with plasmid genomic library construction of *S. pneumoniae* DNA were attributable to multiple effects and not specifically to promoter-like sequences. In addition to the sequence-related effects, cloning of lethal elements and vector-specific problems (copy number) were also cited as possible contributors to the problem. Since *H. pylori* has a lower G + C content (35.2 mol%; Béji *et al.*, 1988) than that of *S. pneumoniae*, sequence-related factors, such as *E. coli* promoter-like sequences, could interfere with genomic library construction. Cloning of lethal elements and plasmid-related effects are also factors which could interfere with *H. pylori* library construction, analogous to the situation with *S. pneumoniae*.

Of all the researchers who constructed genomic libraries of *H. pylori*, only two groups used plasmid vectors. Plasmid libraries were used to clone the *cag*A (Covacci *et al.*, 1993) and *vac*A (Telford *et al.*, 1994) genes. These groups used pBluescript (Stratagene), a phagemid derivative of pUC19, and screened 70,000-85,000 recombinants to obtain their clones. Even with this number of clones neither the *cag*A nor the *vac*A were cloned as intact genes but were cloned as overlapping

fragments. These genes probably represented elements which were toxic to the host strains and thus could not be cloned as entire genes.

Some of the genes cloned for this research (rRNA genes) have homologs in the host strains. In the case where the host chromosome might hybridize to the probe used for screening, plasmid isolations of each individual clone would be necessary. Thus, use of screening procedures such as those used by Covacci *et al.* (1993) and Telford *et al.* (1994) would have been impractical for this work, as it was not feasible to isolate and screen plasmid DNA from 70,000-85,000 clones. Other studies in which plasmid vectors were used in the cloning of *H. pylori* genes involved cloning of specific restriction fragments from genomic digests (Leying *et al.*, 1992; O'Toole *et al.*, 1991; Smith *et al.*, 1994).

Efficient genomic DNA libraries of *H. pylori* could not be produced for this research using the plasmid vector pUC20. Restriction modification was one of the barriers which had to be overcome in the construction of plasmid libraries of *H. pylori* DNA, as indicated by the increased number of recombinants obtained using the DH5αMCR host strain. Other factors, such as those associated with sequence-specific interference (promoter-like sequences), cloning of lethal elements, and vector related problems (copy number), may also have contributed to difficulties observed in plasmid library construction, similar to problems observed by other researchers (Chen and Morrison, 1987; Covacci *et al.*, 1993;

Telford *et al.*, 1994). It appears, therefore, that plasmid vectors may not be well suited for genomic library construction of *H. pylori* DNA.

D.2. Cosmid library construction

When it became apparent that plasmid libraries were not adequate for the purposes of this research, the construction of cosmid libraries was begun in 1990. Since Clayton *et al.* (1989a) had successfully used pHC79 to clone urease genes from *H. pylori*, and the cosmid pSa747 was successfully used for the cloning of *Azotobacter vinelandii* nitrogen fixation genes (Hiratsuka, 1987), these two vectors were chosen for library construction. Unfortunately, an efficient cosmid library of *H. pylori* DNA was not constructed because, as with the plasmid libraries, only a small percentage of the clones carried *H. pylori* DNA. This was probably due to the *E. coli* host strain used for growth and propagation of the cosmid libraries. The *E. coli* strain used was LE392. While the LE392 strain has been mutated for some of the restriction modification functions (*mcr*A$^-$ *hsd*R$^-$), it still retains some restriction modification properties (*mcr*BC$^+$ *hsd*MS$^+$ *mrr*$^+$)(Table 3-4). As mentioned, mutations at the *mcr*C-*hsd*RMS-*mrr* loci are critical for efficient library construction of *H. pylori* DNA. In fact, Phadnis *et al.*(1993) found LE392 to be the least efficient strain for construction of lambda libraries. The transfection efficiencies of the cosmid libraries in an (*mcr*C-*hsd*RMS-*mrr*)$^-$ background were not tested for this

research. This was because, by this time, I had constructed *H. pylori* genomic libraries using the lambda vector λDASH II, and therefore, further work on cloning was concentrated on the lambda libraries. The cosmid pHC79 has been used successfully by a number of groups (Clayton *et al.*, 1989a; Spiegelhalder *et al.*, 1993; Moore *et al.*, 1995). All of these groups used the *E. coli* strain HB101 for growth and propagation of their libraries. The strain HB101 carries a (*mcr*C-*hsd*RMS-*mrr*)$^-$ mutation while retaining the *mcr*A$^+$ phenotype.

While it was possible that the efficiencies of the cosmid libraries might have been greater if they had been transfected into a (*mcr*C-*hsd*RMS-*mrr*)$^-$ host, other factors may also have contributed to the low number of recombinants observed for the cosmid libraries. Factors mentioned previously which could affect the construction of plasmid libraries (low G + C effects, lethal alleles) would be magnified in cosmid clones due to the size of the inserted DNA. Evidence to support this hypothesis was obtained during the construction of a physical genetic map of *H. pylori* using an ordered cosmid library (Bukanov and Berg, 1994). This library was successfully constructed using the cosmid vector Lorist6 in the strain DH5α. This host strain carried a mutation in only the *hsd*M loci and was not mutated for any other restriction modification function (Table 3-4). However, the vector was modified by the addition of rho-dependent transcription terminators bracketing the cloning

site. Such modifications have been used to counteract the instability of cloned DNA associated with the presence of promoter or promoter-like sequences (Chen and Morrison, 1987). In addition, the replication of this vector was regulated by the phage λ replication system. Under this system the copy number of the recombinant cosmids was far more stable that of the standard ColE1 replication system. Using this vector Bukanov and Berg were able to construct a genomic library of *H. pylori* NCTC11638.

The results described here supported observations that LE392 was not an efficient host for *H. pylori* DNA library construction (Phadnis *et al.*, 1993). Results of Bukanov and Berg (1994) indicated that genomic libraries of *H. pylori* DNA could be constructed independent of restriction modification using the vector Lorist6. Problems associated with lethal alleles would still be present. However, construction of more than one library using a variety of enzymes should overcome these problems. In addition, use of a (*mcr*C-*hsd*RMS-*mrr*)$^-$ host may improve library efficiency.

D.3. Lambda library construction

The construction of lambda genomic libraries of *H. pylori* DNA was begun when difficulties with the cosmid libraries were encountered. At this time a restriction map of *H. pylori* UA802 had been constructed in this lab (Taylor *et al.*, 1992a), so efforts to construct genomic libraries of *H. pylori* were concentrated on this strain. The efficient construction of lambda libraries of *H. pylori* genomic DNA

proved to be difficult. No useful *H. pylori* DNA libraries were constructed using λ-GEM11 in the host KW251. Genomic libraries of *H. pylori* UA802 DNA were successfully constructed using the vector λ-DASH II in the host strain SRB, but with low efficiencies. These results were consistent with the later findings of Phadnis *et al.* (1993) that the *E. coli* strains KW251 and SRB were relatively poor hosts for genomic library construction of *H. pylori* DNA. Phadnis *et al.* obtained 2 plaques/μl of phage suspension using *E. coli* KW251 as the host. Similarly, using *E. coli* KW251 as the host for the *H. pylori* UA802 libraries, three plaques/μl of phage suspension (an average of 155.5 plaques/50 μl) were obtained. With *E. coli* SRB, Phadnis *et al.* obtained 4 plaques/μl of phage suspension. For my research, using the host *E. coli* SRB, 124 plaques/μl of phage suspension were obtained (an average of 372 plaques/3 μl). The differences observed in the efficiency of the SRB host may be a reflection of the different lambda vectors used. Phadnis *et al.* used the vector λ-GEM11 (Promega), while I used λ-DASH II (Stratagene).

Other researchers have also reported difficulties in construction of lambda libraries of *H. pylori* DNA. Clayton *et al.* (1989a) reported difficulty using another lambda expression vector λEMBL3. Although sufficient numbers of recombinants were screened, the libraries they constructed were not representative of the entire *H. pylori* genome. Similar findings were made by Covacci *et al.* in 1993, in

their attempt to clone the *cag*A gene using the vectors λ-gt11, λ-EMBL3 and λ-DASH. This gene could not be cloned using λ-EMBL3 and λ-DASH and only portions of the *cag*A were cloned using λ-gt11. Another example of a library which was not representative of the entire genome was that used in the construction of the ordered cosmid map of *H. pylori* NCTC11638 (Bukanov and Berg, 1994). This map contained three gaps for which cosmid recombinants could not be isolated. In addition, there were three sequences which were poorly represented in the cosmid library, one of which was *cag*A, which had been shown by earlier studies to be difficult to clone (Covacci *et al.*, 1993; Tummuru *et al.*, 1993). This suggested that certain regions may harbor genetic elements lethal to *E. coli* or may be regions where the effects of *E. coli* promoter-like sequences were of sufficient strength to interfere with clone stability.

E. Conclusion

Using the lambda vector λ-DASH II, genomic libraries of *H. pylori* UA802 DNA were constructed. These libraries will serve as a source of cloned *H. pylori* for the isolation and characterization of genes from this organism. The construction of genomic libraries of *H. pylori* DNA was also attempted using other vectors. The results of these attempts indicated that restriction modification was a factor that needs to be addressed for the production of genomic libraries of *H. pylori* DNA.

Chapter 4

Isolation and characterization of the P-type ATPase and rRNA genes from the genomic library of *H. pylori* strain UA802

A. Introduction

A.1. P-type ATPase of *H. pylori*

As mentioned in Chapter 1, several substituted benzimidazole compounds have been used in the treatment of peptic and duodenal ulcers. These compounds block the sulfhydral-groups of gastric P-type ATPases and act as proton pump inhibitors to reduce the severity of ulceration (Fellinius *et al.*, 1981; Lindberg *et al.*, 1986; Iwahi *et al.*, 1991). The bacteriostatic, and in some instances bactericidal, activity of these compounds on *H. pylori* suggested the presence of a P-type ATPase. Described in this chapter are the results of research showing the existence of a P-type ATPase in *H. pylori* and the cloning of the genes encoding this ATPase from genomic libraries of *H. pylori* UA802 DNA.

A.2. The 23S and 5S genes of *H. pylori* UA802

Results for the isolation of rRNA genes and characterization of the 23S and 5S genes of *H. pylori* UA802 are also presented in this chapter. Sequence data for the complete 16S rRNA genes of several *H. pylori* strains have already been published (Eckloff *et al.*, 1994). The 23S and 5S genes of *H. pylori* have not been extensively

characterized. The phylogenetic relationships of a variety of *Campylobacter* spp. and related organisms, including *H. pylori*, were investigated by Vandamme *et al.* (1991). In that study, DNA-rRNA hybridizations using radioactively-labelled 23S rRNA molecules were carried out. Vandamme *et al.* gave indirect evidence that *H. pylori* was included in the epsilon subdivision of the class *Proteobacteria*, along with *Campylobacter* spp., *Arcobacter* spp., *Wolinella* spp, *Flexispira* spp. and other *Helicobacter* spp. The phylogenetic position of *H. pylori* was further investigated in this research through the characterization of the 23S and 5S genes.

B. Materials and Methods

Please refer to Chapter 2 for detailed methodology.

C. Results

C.1. Cloning of a P-type ATPase of *H. pylori* UA802

Evidence for the existence of a P-type ATPase in *H. pylori*

Before cloning of genes encoding the P-type ATPase of *H. pylori* was begun, preliminary hybridizations were carried out to show that such genes existed in the genome of *H. pylori*. DNA from *H. pylori* strains UA802 and UA763, digested with *Bgl*II, was fractionated on agarose gels and transferred to nitrocellulose as described (Chapter 2, B.2). The membranes were probed with cloned radioactively-labelled

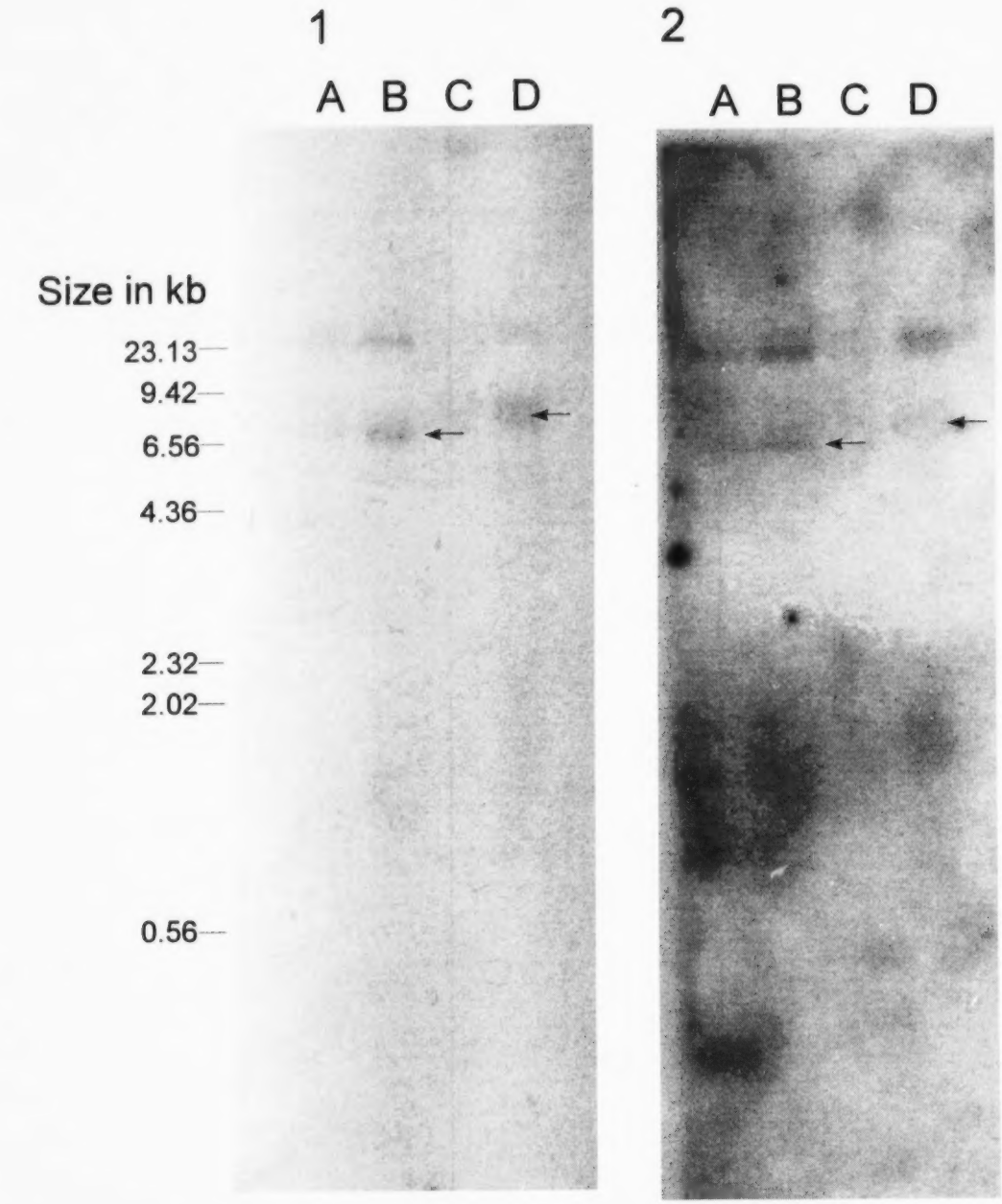
P-type ATPases from three heterologous sources (Figure 4-1). The cloned genes used as probes were the *E. coli kdp*ABC genes (Hesse *et al.*, 1984), the *Enterococcus hirae* (formerly *Streptococcus faecalis*) K^+-ATPase gene (Solioz *et al.*, 1987), and a rat stomach (H^+-K^+)-ATPase gene (Shull and Lingrel, 1986). Figure 4-1. shows Southern analysis of two strains of *H. pylori*, UA763 and UA802, probed with the cloned *E. coli kdp*B gene and *E. hirae* K^+-ATPase gene. Hybridization was very weak and X-ray film had to be exposed to the probed blots for 2 weeks to obtain the results shown in Figure 4-1. Similar results were obtained with the rat (H^++K^+)-ATPase probe (data not shown). These results indicated that both *H. pylori* UA802 and UA763 encoded P-type ATPAses in their genomes.

Screening of the λ-DASH genomic libraries of *H. pylori* UA802 with P-type ATPase probes

Based on the hybridization data, the presence of genes encoding a P-type ATPase in *H. pylori* was established. In order to isolate these genes, PCR probes were produced for screening of the λ-DASH genomic libraries of *H. pylori* UA802 DNA. From amino acid sequence alignments of several P-type ATPases, some regions were observed to be highly conserved (Figure 4-2). Primers for PCR were designed from two of these regions. The primers are shown in Figure 4-2. These primers were used to amplify a PCR probe from *H. pylori* UA802 genomic DNA. A PCR product of approximately 600 bp was amplified from genomic DNA of *H. pylori* UA802

Figure 4-1. Hybridization of restriction digestions of *H. pylori* strains UA763 and UA802 with cloned P-type ATPases

Panel 1 shows hybridization using the cloned *E. hirae* K^+-ATPase probe. Panel 2 shows hybridization using the cloned *E. coli kdp*B probe. Lanes A and B in both panels were 1 μg and 5 μg respectively of *H. pylori* UA802 digested with *Bgl*II. Lanes C and D in both panels were 1 μg and 5 μg respectively of *H. pylori* UA763 digested with *Bgl*II. Arrows indicate hybridizing fragments. Larger hybridizing bands were undigested chromosomal DNA. Both probes hybridized to a fragment of approximately 7 kb in *H. pylori* UA802 and a fragment of approximately 7.5 kb in *H. pylori* UA763. The probes were labelled with [α^{32}P]-dATP. Hybridizations were carried out at 37°C in the presence of 50% formamide.


1
A B C D
2
A B C D
Size in kb
23.13
9.42
6.56
4.36
2.32
2.02
0.56

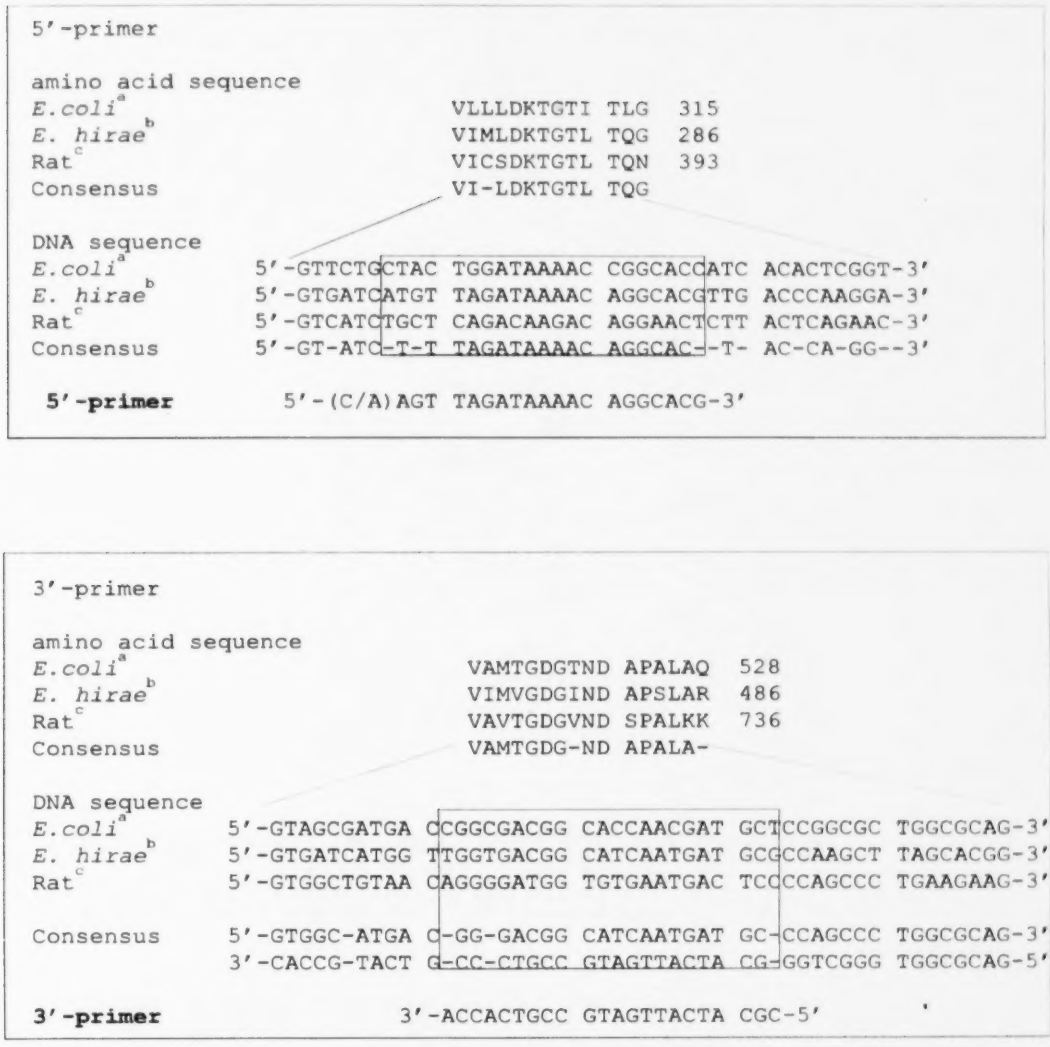

5'-primer
amino acid sequence
E.coli[a] VLLLDKTGTI TLG 315
E. hirae[b] VIMLDKTGTL TQG 286
Rat[c] VICSDKTGTL TQN 393
Consensus VI-LDKTGTL TQG
DNA sequence
E.coli[a] 5'-GTTCTGCTAC TGGATAAAAC CGGCACCATC ACACTCGGT-3'
E. hirae[b] 5'-GTGATCATGT TAGATAAAAC AGGCACGTTG ACCCAAGGA-3'
Rat[c] 5'-GTCATCTGCT CAGACAAGAC AGGAACTCTT ACTCAGAAC-3'
Consensus 5'-GT-ATC-T-T TAGATAAAAC AGGCAC--T- AC-CA-GG--3'
5'-primer 5'-(C/A)AGT TAGATAAAAC AGGCACG-3'
3'-primer
amino acid sequence
E.coli[a] VAMTGDGTND APALAQ 528
E. hirae[b] VIMVGDGIND APSLAR 486
Rat[c] VAVTGDGVND SPALKK 736
Consensus VAMTGDG-ND APALA-
DNA sequence
E.coli[a] 5'-GTAGCGATGA CCGGCGACGG CACCAACGAT GCTCCGGCGC TGGCGCAG-3'
E. hirae[b] 5'-GTGATCATGG TTGGTGACGG CATCAATGAT GCGCCAAGCT TAGCACGG-3'
Rat[c] 5'-GTGGCTGTAA CAGGGGATGG TGTGAATGAC TCCCCAGCCC TGAAGAAG-3'
Consensus 5'-GTGGC-ATGA C-GG-GACGG CATCAATGAT GC-CCAGCCC TGGCGCAG-3'
3'-CACCG-TACT G-CC-CTGCC GTAGTTACTA CG-GGTCGGG TGGCGCAG-5'
3'-primer 3'-ACCACTGCC GTAGTTACTA CGC-5'

Figure 4-2. Regions of conserved amino acid sequence in P-type ATPases used for designing of PCR primers

The top panel shows the region of conserved amino acid sequence used to design the 5' primer. The bottom panel shows the conserved region used to design the 3' primer.
a- from the *Escherichia coli kdp*B sequence, Hesse *et al.*, 1984
b- from the *Enterococcus hirae* K^+-ATPase sequence, Solioz *et al.*, 1987
c- from the rat stomach (H^++K^+)-ATPase sequence, Shull and Lingrel, 1986

(Figure 4-3). Approximately 2000 plaques from each of the λ-DASH *H. pylori* UA802 genomic DNA *Bgl*II and *Sau*3AI libraries were screened using both the cloned *E. hirae* K^+-ATPase gene and the 600 bp PCR product amplified from *H. pylori* UA802. No positive plaques were observed using the *E. hirae* probe. Twenty-three positive plaques, 10 from the *Bgl*II library and 13 from the *Sau*3AI library, were observed using the PCR probe. DNA isolation was carried out on two of these positive plaques, designated λP9 (*Sau*3AI library) and λP14 (*Bgl*II library). Restriction endonuclease digestion analyses were carried out on these clones and the regions containing the P-type ATPase sequences were localized to a 1.3 kb *Eco*RI-*Hind*III fragment of clone λP9 and a 2.7 kb *Eco*RI-*Hind*III fragment of clone λP14 (Figure 4-4).

C.2. PCR amplification of rRNA probes and cloning of the rRNA genes of *H. pylori* UA802

The rRNA PCR probes were amplified from *H. pylori* UA802 using the primers listed in Table 4-1. For the 16S primers, primer 16.5 in combination with primer 16.3a gave a product of approximately 700 bp (Figure 4-5). The combination of primer 16.5 and primer 16.3b gave an almost full-length 16S product of approximately 1500 kb (Figure 4-5).

Nearly full-length 23S PCR products of approximately 2700 bp were amplified using primer 23.5a and primer 23.3a (Figure 4-5), while a product of approximately 650 bp was amplified using the 23.5b primer and 23.3b primer (Figure 4-5). Other combinations of primers (23.5a-23.3b, 23.5b-23.3a) gave

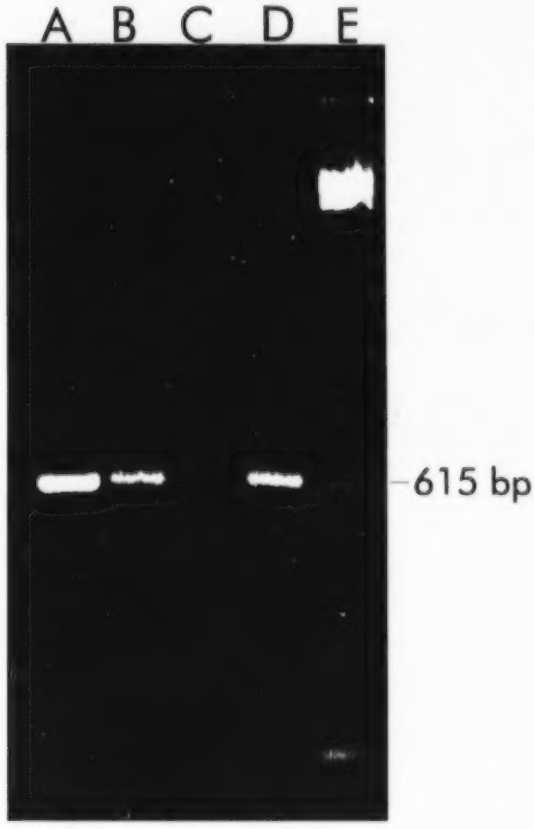


Figure 4-3. Agarose gel analysis of P-type ATPase PCR probe amplified from *H. pylori* UA802

Lanes A-D were 5 µl samples of four separate PCR amplifications of *H. pylori* UA802 using the P-type ATPase primers. Lane E shows the 123 bp ladder size marker. The products were amplified from *H. pylori* UA802 chromosomal DNA using the primers shown on Figure 4-2 in standard PCR reaction conditions (Chapter 2) for 40 cycles of 94°C for 40s, 45°C for 40s and 72°C for 90s. The PCR product in lane C appears to be degraded or did not amplify. The variations in intensities of the other bands reflect the minor variations in template, primer and buffer concentrations which normally occur in the preparation of PCR reactions. The gel used was a 1% agarose gel run in 0.5 X TBE-buffer, set at a constant voltage of 100v.

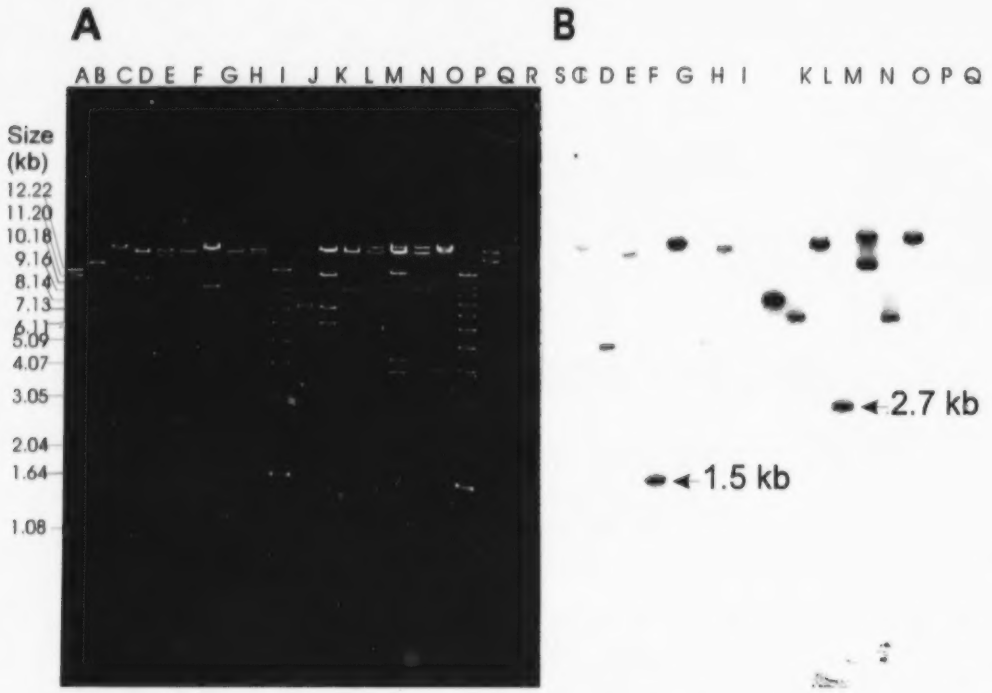


Figure 4-4. Restriction endonuclease digestion analysis and hybridization of λP9 and λP14

Panel A shows the restriction digestions of λP9 and λP14 fractionated on a 0.7% agarose gel; Panel B shows the Southern blot of the gel pictured in panel A, probed with the P-type ATPase PCR probe labelled with [α^{32}P]-dATP. Hybridization was carried out at 37°C in the presence of 50% formamide.

A- λ *Acc*I size std.
B- λ *Ava*I size std.
C- λP9 *Bgl*II
D- λP9 *Bgl*II-*Eco*RI
E- λP9 *Eco*RI
F- λP9 *Eco*RI-*Hind*III
G- λP9 *Hind*III
H- λP9 *Xba*I
I- λP9 *Sal*I
J- 1 kb ladder
K- λP14 *Bgl*II
L- λP14 *Bgl*II-*Eco*RI
M- λP14 *Eco*RI
N- λP14 *Eco*RI-*Hind*III
O- λP14 *Hind*III
P- λP14 *Xba*I
Q- λP14 *Sal*I
R- 1 kb ladder
S- λ *Nco*I size std.
T- λ *Xho*I size std.

Arrows show the 1.5 kb λP9 *Eco*RI-*Hind*III fragment and the 2.7 kb λP14 *Eco*RI-*Hind*III fragment. The 2.7 kb λP14 *Eco*RI-*Hind*III fragment was subsequently subcloned and characterized by Ge *et al.*, 1995.

Table 4-1. Primers used for PCR amplification of 16S and 23S rRNA probes

Gene	Primer sequence	Reference
16S	5'-primer 16.5 5'-GTCGGGTAAGTTCCGACCTG-3' 3'-primers 16.3a 5'-CCGCTTGTCGGTATGGGAAC-3' 16.3b 5'-CCAGTCGCTGTGTG TGCCGTGGGCAGTAGC-3'	Paster *et al.*, 1991
23S	5'-primers 23.5a 5'-GGTGGATGCCTTGGCA(A/C)T-3' 23.5b 5'-GTCGGGTAAGTTCCGACCTG-3' 3'-primers 23.3a 5'-GCTTAGATGCTTTCAGC-3' 23.3b 5'-GGCGAACAGCCATACCCTTG-3'	Egebjerg *et al.*, 1991

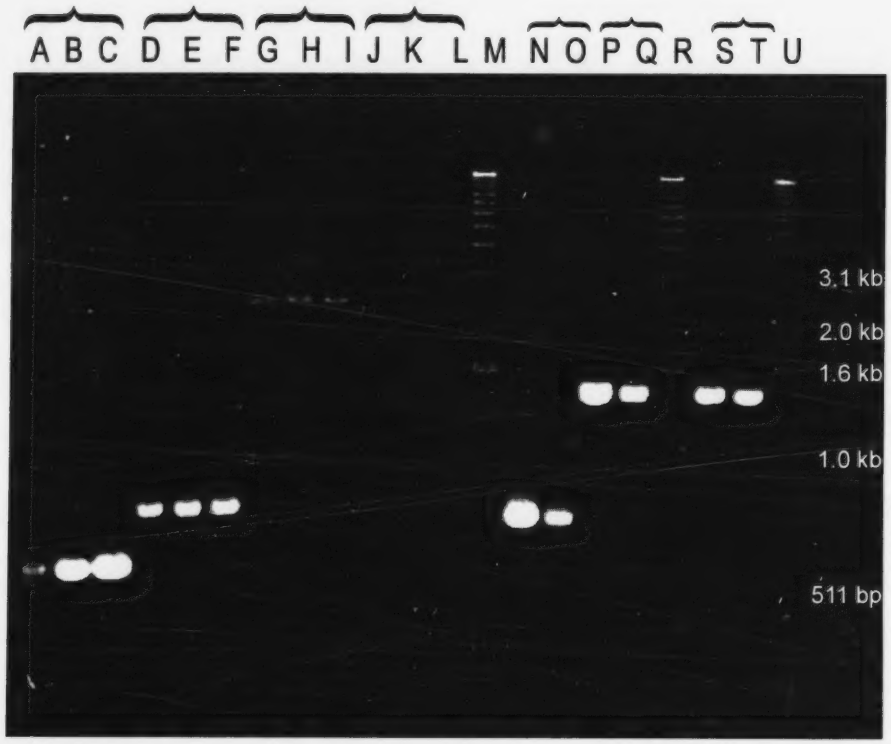


Figure 4-5. Agarose gel analyses of rRNA probes amplified from *H. pylori* UA802 using PCR. Five μl samples of each amplification reaction were run on a 1% agarose gel in 0.5 X TBE-buffer.

Lanes A-C: product amplified with primers 23.5b-23.3b
Lanes D-F: product amplified with primers 23.5b-23.3a
Lanes H-I: product amplified with primers 23.5a-23.3b
Lanes J-L: product amplified with primers 23.5a-23.3a
Lanes M,R,U: 1 kb ladder
Lanes N,O: product amplified with primers 16.5-16.3a
Lanes P,Q,S,T: product amplified with primers 16.5-16.3b.

The products shown in lane J-L represent almost the entire 23S gene (full-length 23S genes are approximately 2900 bp), while the products in lanes P, Q, S, and T represent almost the entire 16S gene (full length 16S genes are approximately 1600 bp).

products of sizes intermediate to the 650 bp and 2700 bp products (Figure 4-5).

Screening of the λ-DASH genomic libraries of *H. pylori* UA802 with rRNA probes

The λ-DASH genomic libraries (*Bgl*II and *Sau*3AI) of *H. pylori* UA802 DNA were screened with the 700 bp 16S-PCR probe for the 16S rRNA genes and the 650 bp 23S-PCR probe for the 23S rRNA genes. Approximately 2000 plaques were screened from each library. With the 16S probe approximately 18% of the clones were positive, while with the 23S probe approximately 7.5% of the clones were positive. Phage stocks were made from 16S and 23S clones randomly chosen from the *Bgl*II and *Sau*3AI libraries. DNA was isolated from ten of the 16S stocks and ten of the 23S stocks. Figure 4-6 shows the restriction analyses of one of the 16S clones (λ16.4) and one of the 23S clones (λ23.54). Low resolution restriction maps of λ16.4 and λ23.54 were made by partial digestion mapping (Figure 4-7). The insert in λ16.4 was found to be appproximately 16.4 kb while the insert in λ23.54 was found to be approximately 20 kb. The 600 bp 16S probe was shown to hybridize to the 1 kb *Eco*RI fragment of λ16.4. The 700 bp 23S probe hybridized to the 10 kb *Xba*I fragment of λ23.54.

Isolation of each individual copy of the 23S genes was carried out using λ23.54 as a starting point. The 23S gene was localized to a 10 kb *Xba*I fragment of this clone (Figures 4-6 and 4-7). The 3 kb *Eco*RI-*Xba*I fragment did not hybridize with the 23S probe and was localized to the opposite end to

Figure 4-6. Agarose gel electrophoresis of *Eco*RI and *Xba*I restriction digestions of clones λ16.4 and λ23.54, and hybridization of Southern transfers of the gels with either the 16S or 23S PCR probes.

Lane A- λ16.4 *Eco*RI

Lane B- λ16.4 *Xba*I

Lane C- lane A hybridized with the 700 bp 16S PCR probe

Lane D- lane B hybridized with the 700 bp 16S PCR probe

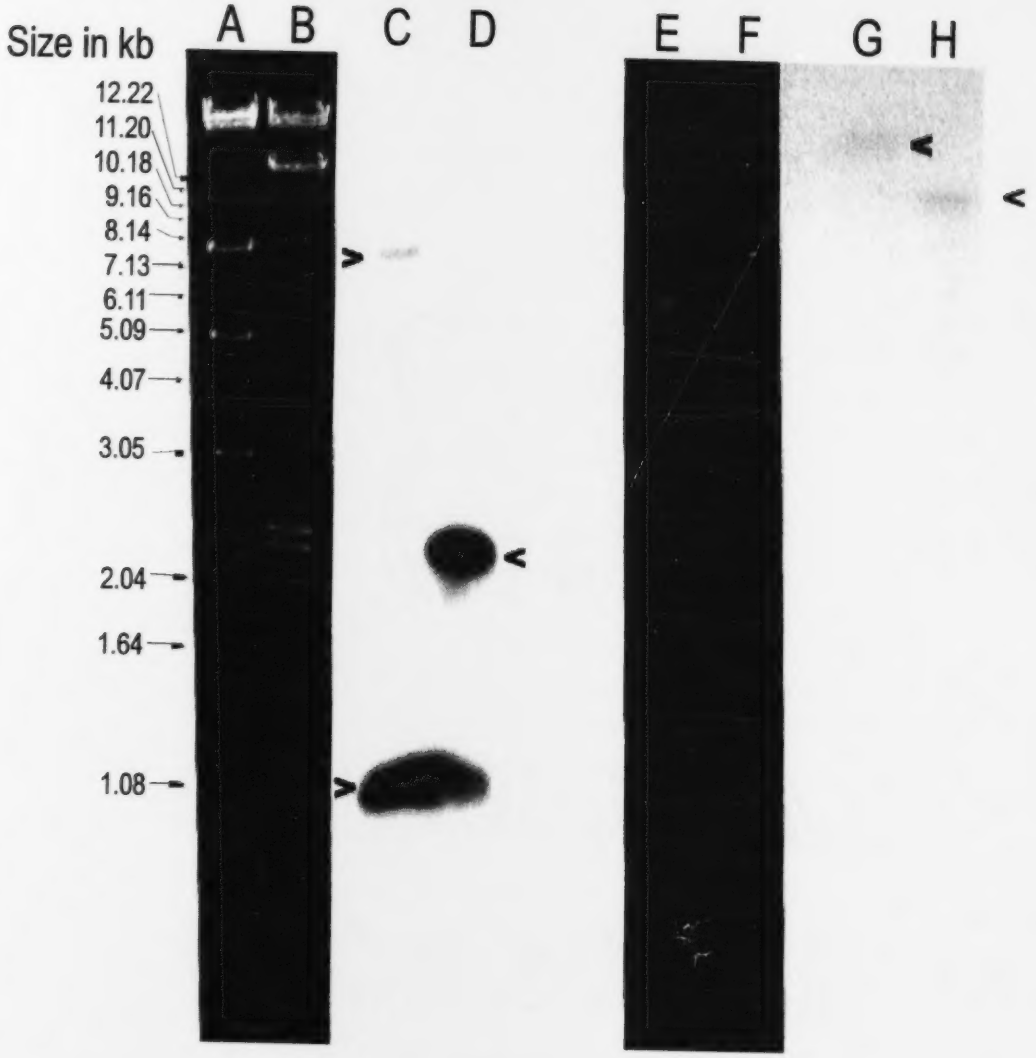
Lane E- λ23.54 *Eco*RI

Lane F- λ23.54 *Xba*I

Lane G- Lane E hybridized with the 650 bp 23S PCR probe

Lane H- Lane F hybridized with the 650 bp 23S PCR probe

The probes were labelled with [α^{32}P]-dATP. Hybridizations were carried out at 37°C in the presence of 50% formamide.


Size in kb
A
B
C
D
E
F
G
H
12.22
11.20
10.18
9.16
8.14
7.13
6.11
5.09
4.07
3.05
2.04
1.64
1.08

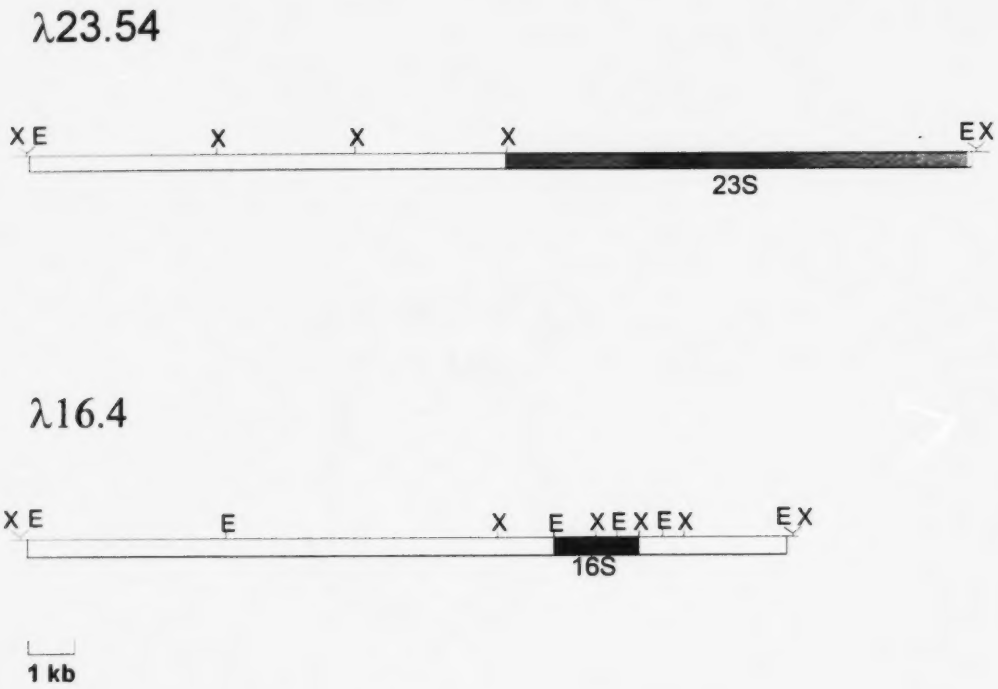


Figure 4-7. Low resolution restriction maps of clones λ16.4 and λ23.54 constructed using partial digestion mapping, showing the location of the 16S and 23S genes

E- *Eco*RI
X- *Xba*I
The rRNA genes were localized using data obtained in the restriction endonuclease digestion analyses of these clones (Figure 4-6) and are indicated by shading.

the 23S gene on the inserted *H. pylori* UA802 DNA fragment of λ23.54. The 3 kb *Eco*RI-*Xba*I fragment was isolated from low melting point agarose and hybridized to restriction digestions of nine 23S DNA λ-clones (Figure 4-8). Clones isolated from the same region of the genome of *H. pylori* UA802 might also carry this fragment. Thus all clones hybridizing to this fragment could be eliminated and clones which did not hybridize could have originated from the other 23S region on the *H. pylori* UA802 genome. By using this hybridization data in combination with comparisons of restriction endonuclease digestion patterns, clones carrying the other copy of the 23S were selected. The lambda vector was expected to give fragments of 9 and 22 kb in the *Xba*I digests. The 3 kb *Eco*RI-*Xba*I fragment appeared to have homology with lambda, as the 9 and 22 kb lambda bands hybridized to the probe (Figure 4-8). However, Figure 4-8 also shows that the inserted *H. pylori* UA802 DNA of clones λ23.B1, λ23.B2, λ23.S1 and λ23.S4 had no homology with the probe, indicating that the clones may have originated from a different region of the genome than λ23.54. Comparisons of restriction endonuclease patterns also suggested that these clones had different chromosomal origins than λ23.54. Large scale DNA isolation was carried out on λ23.B1 for further characterization. No further characterization was done with clones λ23.B2, λ23.S1 and λ23.S4. With clones λ23.54 and λ23.B1, two copies of the 23S genes of *H. pylori* UA802 were

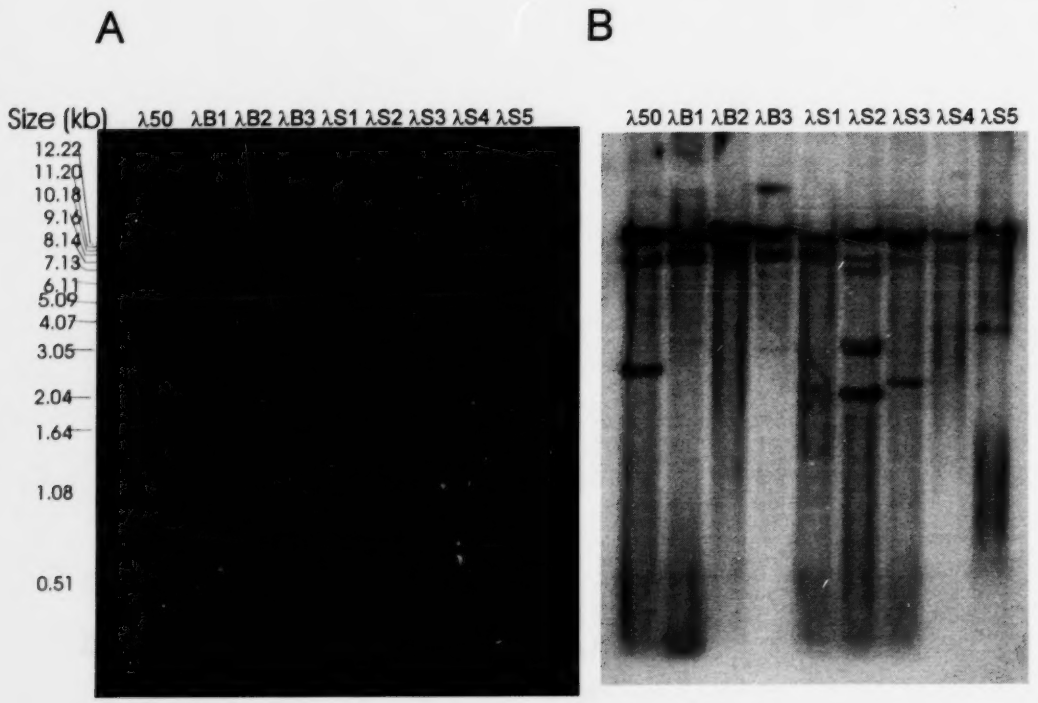


Figure 4-8. Agarose gel electrophoresis of *Xba*I digestions of nine 23S λ-clones and hybridization of these clones with DNA from the region flanking the 23S gene of λ23.54

Panel A shows *Xba*I digestions of 23S lambda clones fractionated on 0.6% agarose. With *Xba*I, the lambda vector gives fragments of 9 kb and 22 kb.

Panel B shows a Southern blot of the gel hybridized with the 2 kb *Eco*RI-*Xba*I fragment of λ23.54. The probe hybridized with the lambda bands and did not hybridize with the insert DNA of λ23.B1, λ23.B2, λ23.S1, and λ23.S4. The probes were labelled with [α^{32}P]-dATP. Hybridizations were carried out at 42°C in the presence of 50% formamide.

represented and the 23S sequences from these clones were subcloned and characterized.

Several attempts to amplify probes for the 5S gene of *H. pylori* UA802 were carried out in our laboratory but were not successful (E. Newnham, personal communication). However, the 5S genes are normally found immediately 3' to the 23S genes in procaryotes (Nomura *et al.*, 1984) and it was likely that the 5S gene would be present, along with the 23S genes, in clones λ23.54 and λ23.B1.

Subcloning of the 23S and 5S rRNA genes

For further characterization of the rRNA genes the DNA fragments harboring the 23S genes were further subcloned into the plasmid pBluescript SK⁻ (Stratagene). The 23S genes of both λ23.54 and λ23.B1 were localized to *Xba*I fragments of approximately 10 kb in size (data not shown). These fragments were cloned into the pBluescript *Xba*I site to give the plasmids p54X1.1 and pB1X1.1. Restriction maps of these clones were made with the enzymes *Kpn*I, *Sst*II and *Nru*I (Figure 4-9). Restriction digestions of these clones were blotted onto nitrocellulose and hybridized with the 700 bp and 2800 bp 23S PCR probes. A 1 kb *Sst*II fragment was located within the 23S genes of both p54X1.1 and pB1X1.1. These 1 kb *Sst*II fragments were further subcloned into pBluescript as were the *Sst*II-*Xba*I fragments flanking the 1 kb fragments. Subcloning of the *Nru*I-*Xba*I fragments was also done.

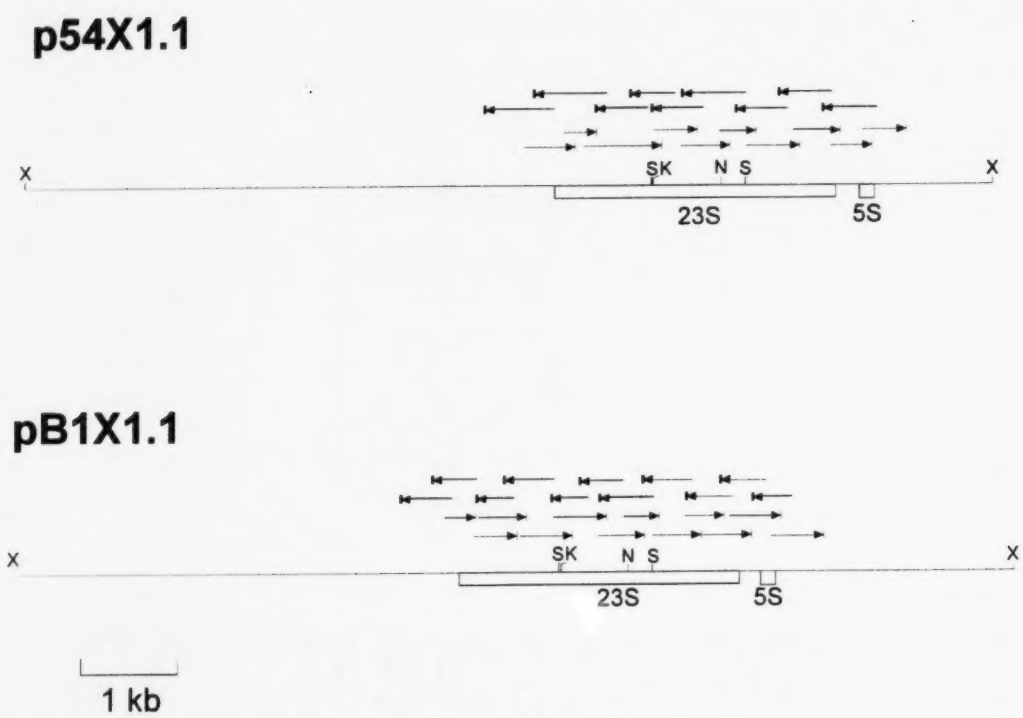


Figure 4-9. Restriction maps of p54X1.1 and pB1X1.1 subcloned from λ23.54 and λ23.B1 respectively, showing the location of the 23S and 5S genes and sequencing strategy used to determine rRNA gene sequences

S- *Sst*II
K- *Kpn*I
N- *Nru*I
X- *Xba*I

p54X1.1 was the cloned 10 kb *Xba*I fragment of λ23.54 and pB1X1.1 was the cloned 10.2 kb *Xba*I fragment of λB1. The figure also shows the sequencing strategy used. Arrows above the map show the length and direction of segments sequenced.

Sequencing of the 23S and 5S genes of *H. pylori* UA802

DNA sequence was obtained from all of the plasmids described above using cycle sequencing and either T7 or T3 sequencing primer. Sequential deletions of some of these subclones were made. Custom sequencing primers were constructed as sequence became available. These primers are listed on Table 4-2 and are shown on Figure 4-10. The sequencing strategy used is shown in Figure 4-9. The complete sequences of the 23S and 5S genes as well as sequence from both the 5'- and 3'-flanking regions and the intervening sequence between the 23S and 5S genes of both clones were determined. Both sequences were identical starting from 373 bp 5' to the 23S gene, through the 228 bp intervening sequence and up to 61 bp beyond the 3' end of the 5S gene, at which point the sequences diverged (Figure 4-10).

Primer extension of 23S and 5S rRNA

Total RNA was extracted from *H. pylori* UA802, and primer extensions were carried out using the primers listed in Table 4-2. From this analysis the starting points of the 23S and 5S rRNA molecules could be determined. The start point of the 23S rRNA molecule was found to be an A nucleotide at position 373 of the sequence (Figures 4-10 and 4-11). Five different start sites were found for the 5S rRNA molecule, four T nucleotides at positions 3568,3572, 3573, and 3574, and one C nucleotide at position 3571 (Figures 4-10 and 4-11).

Table 4-2. Primers used for sequencing and primer extension of the 23S and 5S genes of *H. pylori* UA802

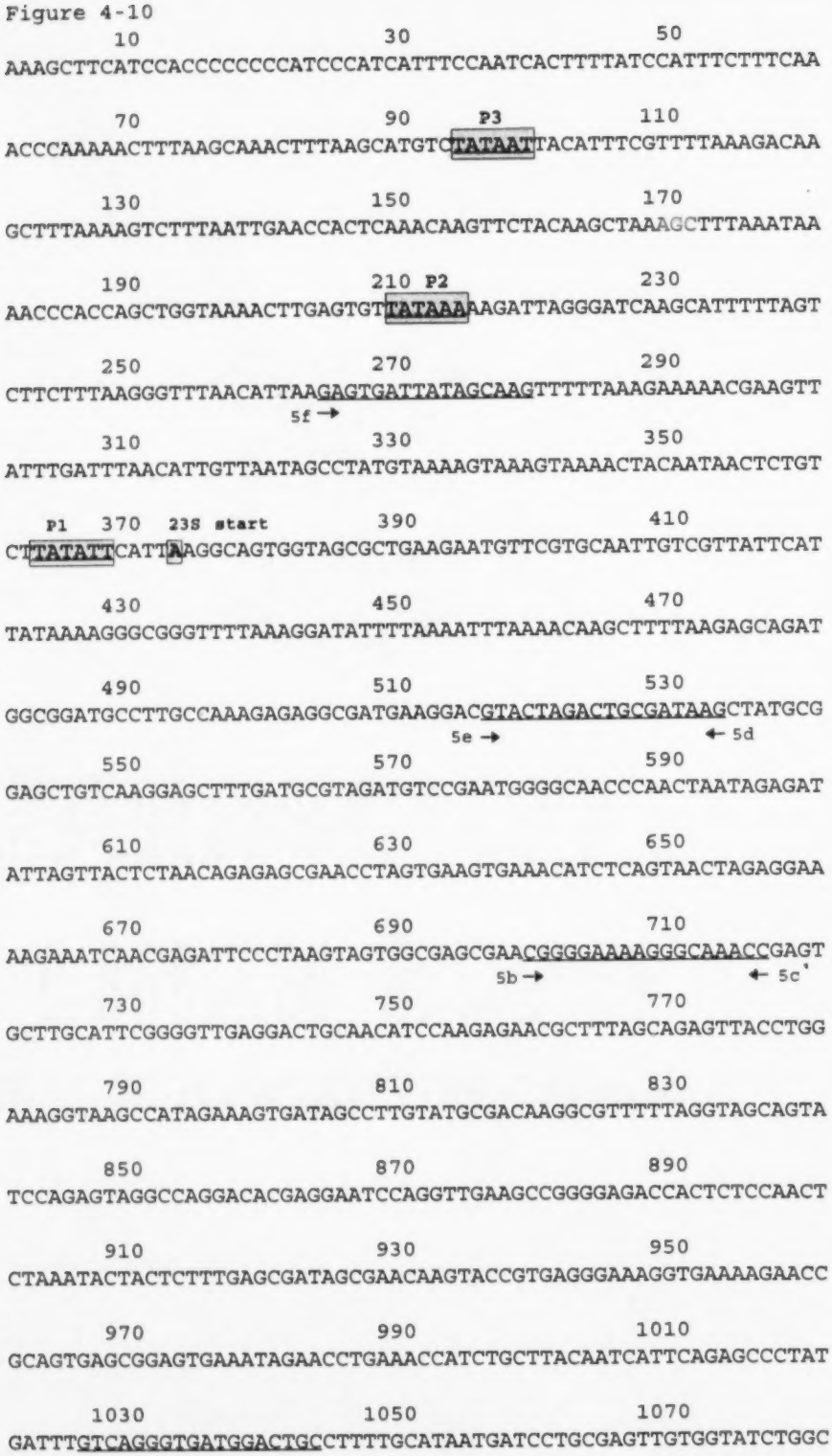
Primer	Application	Sequence (5'-3')	Position of 5'-end*
5b	sequencing	CGGGGAAAAGGGCAAACC	base 699 (s)
5c	sequencing	GGTTTGCCCTTTTCCCCG	base 716 (as)
5d	sequencing/ primer extension	CTTATCGCAGTCTAGTAC	base 533 (as)
5e	sequencing	GTACTAGACTGCGATAAG	base 516 (s)
5f	sequencing	GAGTGATTATAGCAAG	base 264 (s)
5g	sequencing	GTCAGGGTGATGGACTG	base 1026 (s)
5h	sequencing	GACTATGTGCTTGATAG	base 3257 (s)
5i	sequencing	CCATTCCGAACCTG	base 3611 (s)
5j	sequencing	GGACTGAACTCCTACCC	base 1207 (s)
3a	sequencing/ PCR	GGCGAACAGCCATACCCTTG	base 3014 (as)
3b	sequencing/ primer extension	GAGCAGTATTATCAG	base 3651 (as)
3c	sequencing	CGACATCGAGGTGCC	base 2963 (as)
3d	sequencing	CCAGCACTCCTTACC	base 1711 (as)
3e	sequencing	CTAGCCCAATCAGCGC	base 1358 (as)

*- position of the first nucleotide of the primer on the sequence shown in Figure 4-10. Whether the primer is sense (s) or anti-sense (as) is indicated in brackets.

Figure 4-10. Complete nucleotide sequences of the 23S and 5S genes as well as part of the flanking region from clones p54X1.1 and pB1X1.1

–Translation signals (Pribnow boxes) are marked by underlining and are highlighted.
–Start positions of the 23S and 5S rRNA molecules as determined by primer extension are in bold and are highlighted.
–Putative end positions are in bold and are highlighted.
–The 3' sequences flanking the rRNA genes are shown up to the point where the sequences of the two clones diverge.
-Primers used for sequencing and primer extension, listed in Table 4-2 are underlined and labelled. The direction of the primers are indicated by arrows.
–The DNA sequence shown in this figure has been submitted to Genbank: accession number U27270.

Figure 4-10


10
30
50
AAAGCTTCATCCACCCCCCCCCATCCCATCATTTCCAATCACTTTTATCCATTTCTTTCAA
70
90
P3
110
ACCCAAAAACTTTAAGCAAACTTTAAGCATGTCTATAATTACATTTCGTTTTAAAGACAA
130
150
170
GCTTTAAAAGTCTTTAATTGAACCACTCAAACAAGTTCTACAAGCTAAAGCTTTAAATAA
190
210
P2
230
AACCCACCAGCTGGTAAAACTTGAGTGTTATAAAAAGATTAGGGATCAAGCATTTTTAGT
250
270
290
CTTCTTTAAGGGTTTAACATTAAGAGTGATTATAGCAAGTTTTTAAAGAAAAACGAAGTT
5f →
310
330
350
ATTTGATTTAACATTGTTAATAGCCTATGTAAAAGTAAAGTAAAACTACAATAACTCTGT
P1
370
23S start
390
410
CTTATATTCATTAAGGCAGTGGTAGCGCTGAAGAATGTTCGTGCAATTGTCGTTATTCAT
430
450
470
TATAAAAGGGCGGGTTTTAAAGGATATTTTAAAATTTAAAACAAGCTTTTAAGAGCAGAT
490
510
530
GGCGGATGCCTTGCCAAAGAGAGGCGATGAAGGACGTACTAGACTGCGATAAGCTATGCG
5e →
← 5d
550
570
590
GAGCTGTCAAGGAGCTTTGATGCGTAGATGTCCGAATGGGGCAACCCAACTAATAGAGAT
610
630
650
ATTAGTTACTCTAACAGAGAGCGAACCTAGTGAAGTGAAACATCTCAGTAACTAGAGGAA
670
690
710
AAGAAATCAACGAGATTCCCTAAGTAGTGGCGAGCGAACGGGGAAAAGGGCAAACCGAGT
5b →
← 5c'
730
750
770
GCTTGCATTCGGGGTTGAGGACTGCAACATCCAAGAGAACGCTTTAGCAGAGTTACCTGG
790
810
830
AAAGGTAAGCCATAGAAAGTGATAGCCTTGTATGCGACAAGGCGTTTTTAGGTAGCAGTA
850
870
890
TCCAGAGTAGGCCAGGACACGAGGAATCCAGGTTGAAGCCGGGGAGACCACTCTCCAACT
910
930
950
CTAAATACTACTCTTTGAGCGATAGCGAACAAGTACCGTGAGGGAAAGGTGAAAAGAACC
970
990
1010
GCAGTGAGCGGAGTGAAATAGAACCTGAAACCATCTGCTTACAATCATTCAGAGCCCTAT
1030
1050
1070
GATTTGTCAGGGTGATGGACTGCCTTTTGCATAATGATCCTGCGAGTTGTGGTATCTGGC

5g→

1090 1110 1130
AAGGTTAAGCGAATGCGAAGCCGTAGCGAAACGAGTTCTTAATAGGGCGAACAAGTCAGA

1150 1170 1190
TGCTGCAGACCCGAAGCTAAGTGATCTATCCATGGCCAAGTTGAAACGCGTGTAATAGCG

Figure 4-10 cont'd

1210 1230 1250
CGTGGAGGACTGAACTCCTACCCATTGAAACGGGTTGGGATGAGCTGTGGATAGGGGTGA

5j→

1270 1290 1310
AAGGCCAAACAAACTTAGTGATAGCTGGTTCTCTTCGAAATATATTTAGGTATAGCCTCA

1330 1350 1370
AGTGATAATAAAAGGGGGTAGAGCGCTGATTGGGCTAGGGCTGCTCGCCGCGGTACCAAA

←3e

1390 1410 1430
CCCTATCAAACTTCGAATACCTTTTATCGTATCTTGGGAGTCAGGCGGTGGGTGATAAAA

1450 1470 1490
TCAATCGTCAAAAGGGGAACAACCCAGACTACCAAATAAGGTCCCTAAGTTCTATTCTGA

1510 1530 1550
GTGGAAAAAGATGTGTGGCTACTAAAACAACCAGGAGGTTGGCTTAGAAGCAGCCATCCT

1570 1590 1610
TTAAAGAAAGCGTAACAGCTCACTGGTCTAGTGGTCATGCGCTGAAAATATAACGGGGCT

1630 1650 1670
AAGATAGACACCGAATTTGTAGATTGTGTTAAACACAGTGGTAGAAGAGCGTTCATACCA

1690 1710 1730
GCGTTGAAGGTATACCGGTAAGGAGTGCTGGAGCGGTATGAAGTGAGCATGCAGGAATGA

←3d

1750 1770 1790
GTAACGATAAGATATATGAGAATTGTATCCGCCGTAAATCTAAGGTTTCCTACGCGATGG

1810 1830 1850
TCGTCATCGTAGGGTTAGTCGGGTCCTAAGCCGAGTCCGAAAGGGGTAGGTGATGGCAAA

1870 1890 1910
TTGGTTAATATTCCAATACCGACTGTGGAGCGTGATGGGGGGACGCATAGGGTTAAGCGA

1930 1950 1970
GCTAGCTGATGGAAGCGCTAGTCTAAGGGCGTAGATTGGAGGGAAGGCAAATCCACCTCT

1990 2010 2030
GTATTTGAAACCCAAACAGGCTCTTTGAGTCCTTTTAGGACAAAGGGAGAATCGCTGATA

2050 2070 2090
CCGTCGTGCCAAGAAAAGCCTCTAAGCATATCCATAGTCGTCCGTACCGCAAACCGACAC

2110 2130 2150
AGGTAGATGAGATGAGTATTCTAAGGCGCGTGAAAGAACTCTGGTTAAGGAACTCTGCAA

2170 2190 2210
ACTAGCACCGTAAGTTCGCGATAAGGTGTGCCACAGCGATGTGGTCTCAGCAAAGAGTCC

2230 2250 2270
CTCCCGACTGTTTACCAAAAACACAGCACTTTGCCAACTCGTAAGAGGAAGTATAAGGTG

2290 2310 2330
TGACGCCTGCCCGGTGCTCGAAGGTTAAGAGGATGCGTCAGTCGCAAGATGAAGCGTTGA

2350 2370 2390
ATTGAAGCCCGAGTAAACGGCGGCCGTAACTATAACGGTCCTAAGGTAGCGAAATTCCTT

Figure 4-10 cont'd

2410 2430 2450
GTCGGTTAAATACCGACCTGCATGAATGGCGTAACGAGATGGGAGCTGTCTCAACCAGAG

2470 2490 2510
ATTCAGTGAAATTGTAGTGGAGGTGAAAATTCCTCCTACCCGCGGCAAGACGGAAAGACC

2530 2550 2570
CCGTGGACCTTTACTACAACTTAGCACTGCTAATGGGAATATCATGCGCAGGATAGGTGG

2590 2610 2630
GAGGCTTTGAAGTAAGGGCTTTGGCTCTTATGGAGTCATCCTTGAGATACCACCCTTGAT

2650 2670 2690
GTTTCTGTTAGCTAACTGGCCTGTGTTATCCACAGGCAGGACAATGCTTGGTGGGTAGTT

2710 2730 2750
TGACTGGGGCGGTCGCTCCTAAAAAGTAACGGAGGCTTGCAAAGGTTGGCTCATTGCGGT

2770 2790 2810
TGGAAATCGCAAGTTGAGTGTAATGGCACAAGCCAGCCTGACTGTAAGACATACAAGTCA

2830 2850 2870
AGCAGAGACGAAAGTCGGTCATAGTGATCCGGTGGTTCTGTGTGGAAGGGCCATCGCTCA

2890 2910 2930
AAGGATAAAAGGTACCCCGGGGATAACAGGCTGATCTCCCCCAAGAGCTCACATCGACGG

2950 2970 2990
GGAGGTTTGGCACCTCGATGTCGGCTCATCGCATCCTGGGGCTGGAGCAGGTCCCAAGGG

←3c

3010 3030 3050
TATGGCTGTTCGCCATTTAAAGCGGTACGCGAGCTGGGTTCAGAACGTCGTGAGACAGTT

←3a

3070 3090 3110
CGGTCCCTATCTGCCGTGGGCGTAGGAAAGTTGAGGAGAGCTGTCCCTAGTACGAGAGGA

3130 3150 3170
CCGGGATGGACGTGTCACTGGTGCACCAGTTGTCTGCCAAGAGCATCGCTGGGTAGCACA

3190 3210 3230
CACGGATGTGATAACTGCTGAAAGCATCTAAGCAGGAACCAACTCCAAGATAAACTTTCC

```
      3250                 3270                 3290
CTGAAGCTCGCACAAAGACTATGTGCTTGATAGGGTAGATGTGTGAGCGCAGTAATGCGT
                5h →
      3310                 3330  23S ends       3350
TTAGCTGACTACTACTAATAGAGCGTTTGGCTTGTTTTTTGCTTTTTGATAAGATAACGG

      3370                 3390                 3410
CAATAAGCGCGAATGGGTTACCACTGCCTTACTGAGTGTAAGAGAGTTGGAGTTTTATGA

      3430                 3450                 3470
AGACTTTTATAAGATTAAACTTTAATGAGGAATGAGATACCATCTCAATGGTTTAAAGTT

      3490                 3510                 3530
AAAGGCTATTAACGATCTTCTTTGTTAAAAACAGCTCCCCTATAAAGAGAAAGGGGAGTT

      3550                 3570 5S begins       3590
AAGGGTAAATGCGTTTTTATCTTTAGCTCCCTTTTCCTTGTGCCTTTAGAGAAGAGGAAC
```

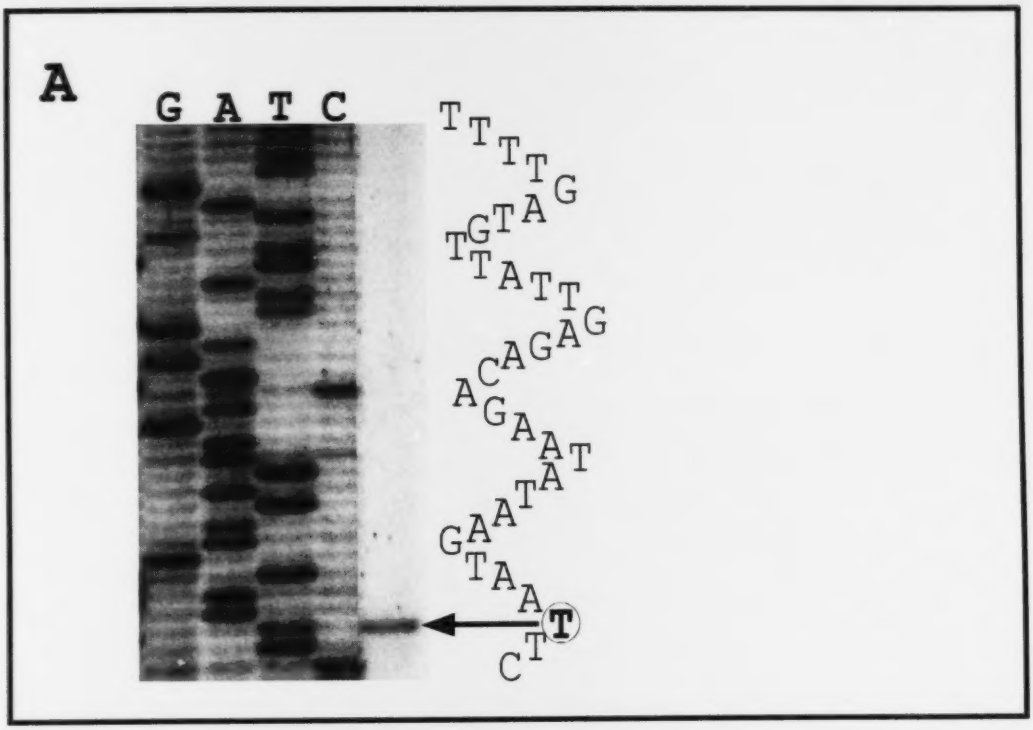
Figure 4-10 cont'd

```
      3610                 3630                 3650
TACCCAGTTAACCATTCCGAACCTGGAAGTCAAGCTCTTCATCGCTGATAATACTGCTCT
        5i →                                              ← 3b
      3670                 3690     5S ends     3710
TTTCAAGAGTGGGAATGTAGGTCGGTGCAGGGATAGGGAAATGTTTTTTTAGTCTTGCTT

      3730                 3750
TTTTATTTGATTTCATTATTGACTCATTGTTTTGTTTGTTTAGG
```

Figure 4-11. Primer extension of total RNA of *H. pylori* UA802 to determine the start points of the 23S and 5S rRNA molecules

Panel A shows primer extension using primer 5d to determine the start position of the 23S rRNA molecule.

Panel B shows primer extension using primer 3b to determine the start position of the 5S rRNA molecule.


A
G A T C
T T T T G
T G T A
T T A T T
A C A G A G
A G A A A T
G A A T A
G T A A
T
C T

B
G A T C
A G A A A
C T A
G C
A
G G
G
A
A
A
A G G A A

PCR from divergent sequence

Because the sequences of the two copies of the 23S/5S genes showed so much identity, it was possible that only one copy of the genes was represented by these two clones. At some point during the cloning of the genes some rearrangement may have occurred to give the appearance that two copies diverging 61 bp from the end of the 5S gene that had been cloned. In order to show that two copies had in fact been cloned, primers were designed from the sequence of the clones within the regions where they were different (Figure 4-12). These primers, 54.ext and B1.ext, were then used to amplify products from the plasmid clones (p54X1.1 and pB1X1.1), the lambda clones (λ23.54 and λB1) and *H. pylori* UA802 chromosomal DNA in combination with the primers 5h and 5i (Table 4-2). The primer 5h starts 82 bp from the end of the 23S gene and primer 5i starts 90 bp from the end of the 5S gene. In combination with primer 5h, a product of 655 bp is expected with primer 54.ext and a product of 569 bp is expected with primer B1.ext. In combination with primer 5i, a product of 300 bp is expected with primer 54.ext and a product of 214 bp is expected with primer B1.ext. If the clones do represent both copies of the 23S/5S genes, then products amplified using the 54.ext primer should only be amplified from p54X1.1, λ23.54 and the *H. pylori* UA802 chromosomal DNA. Likewise, using the primer B1.ext, products should only be amplified from pB1X1.1, λB1 and the *H. pylori* UA802 chromosomal DNA. Rearranged clones would either have a

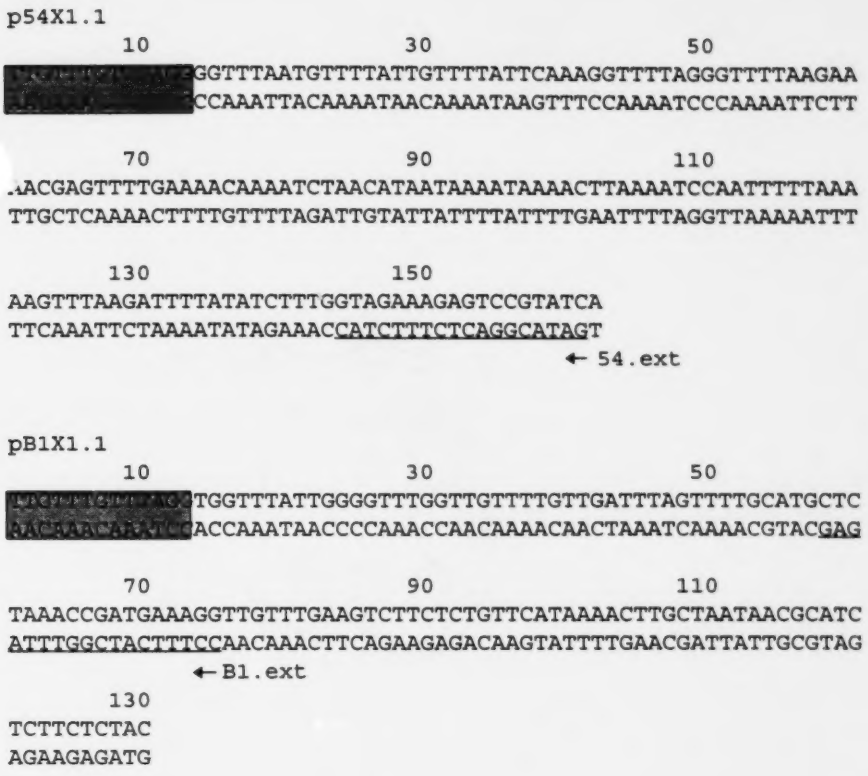


Figure 4-12. Sequence of DNA from the points where p54X1.1 and pB1X1.1 diverge.

Shaded boxes show end of sequence up to where the sequences of the two clones were identical. The last nucleotide in the boxes corresponds to the last nucleotide of the sequence shown in Figure 4-10. Underlined sequences indicate sequence used as PCR primers and the direction of these primers is indicated with an arrow.

different-sized product or no product amplified from the *H. pylori* UA802 chromosomal DNA. Figure 4-13 shows that for all primer combinations, the expected products were amplified from the specific clones and from the *H. pylori* UA802 chromosomal DNA.

Phylogenetic comparisons using the 23S and 5S genes of *H. pylori*

Sequence analyses and phylogenetic comparisons were done using the Wisconsin Sequence Analysis Package (GCG) as outlined in Chapter 2. The complete sequence of the 23S gene of *H. pylori* UA802 was compared to the complete sequences of the 23S genes of *C. jejuni*, *C. coli*, *Pseudomonas cepacia* (accession #X16368, Hopfl *et al.*, 1989), *E. coli* (accession #V00331, Brosius *et al.*, 1980), *Rhodobacter capsulatus* (accession #X06485, Hopfl *et al.*, 1988), *Anacysistis nidulans* (accession #X00512, Douglas and Doolittle, 1984), *Bacillus subtilis* (accession #D11460, Wawrousek and Hansen, 1983) and *Micrococcus luteus* (accession #X06484, Regensburger *et al.*, 1988) using the pileup program of the Wisconsin Sequence Analysis Package. Figure 4-14 shows the dendogram constructed from this comparison and represents the graphical output of the pileup program.

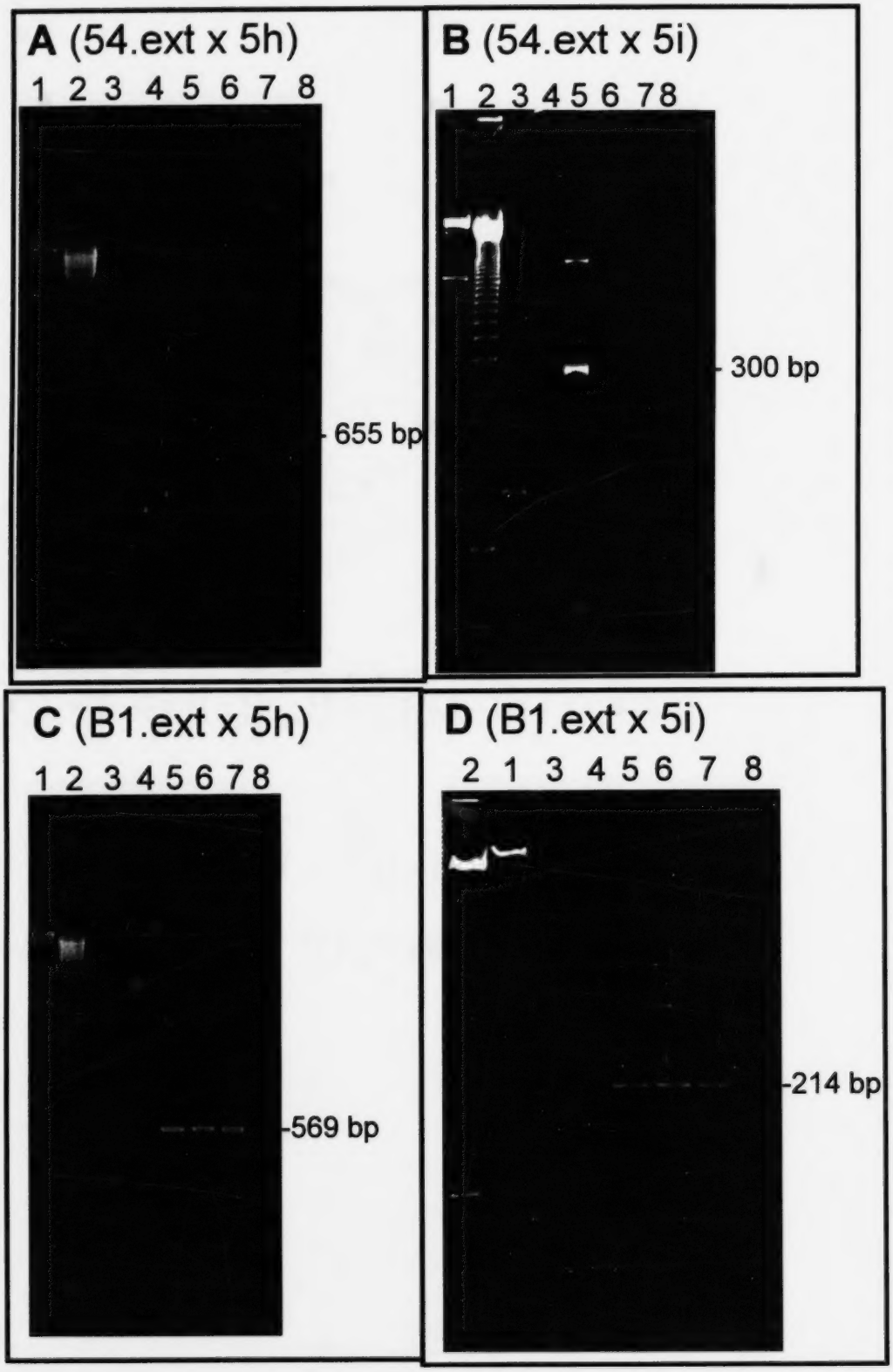
Phylogenetic comparisons were also carried out with the partial 23S sequences of several *Campylobacter* species (Figure 4-15)(van Camp *et al.*, 1993), again using the pileup program of GCG. For these comparisons, only the regions of the *H. pylori* UA802, *C. jejuni* (Kim *et al.*, 1994) and *C. coli*

Figure 4-13. PCR amplification from p54X1.1, pB1X1.1, λ23.54, λB1 and *H pylori* UA802 chromosomal DNA using primers designed from sequences 3' to the 5S gene where the p54X1.1 and pB1X1.1 sequences differ.

For all panels:
1- 1 kb ladder (BRL)
2- 123 bp ladder (BRL)
3- product amplified from p54X1.1
4- product amplified from λ23.54
5- product amplified from *H pylori* UA802 chromosomal DNA
6- product amplified from pB1X1.1
7- product amplified from λB1
8- no DNA; negative control

Panel A- primers 54.ext and 5h
Panel B- primers 54.ext and 5i
Panel C- primers B1.ext and 5h
Panel D- primers B1.ext and 5i

Panels A and C show 5μl aliquots of the PCR products electrophoresed through 1.5% agarose gels which were run with 100v constant voltage in 0.5 X and then stained with ethidium bromide.
Panels B and D show 5μl aliquots of the PCR products electrophoresed through 5% acrylamide gels which were run with 300v constant voltage in 0.5 X and then stained with ethidium bromide.


A (54.ext x 5h)
1 2 3 4 5 6 7 8
655 bp
B (54.ext x 5i)
1 2 3 4 5 6 7 8
300 bp
C (B1.ext x 5h)
1 2 3 4 5 6 7 8
569 bp
D (B1.ext x 5i)
2 1 3 4 5 6 7 8
214 bp

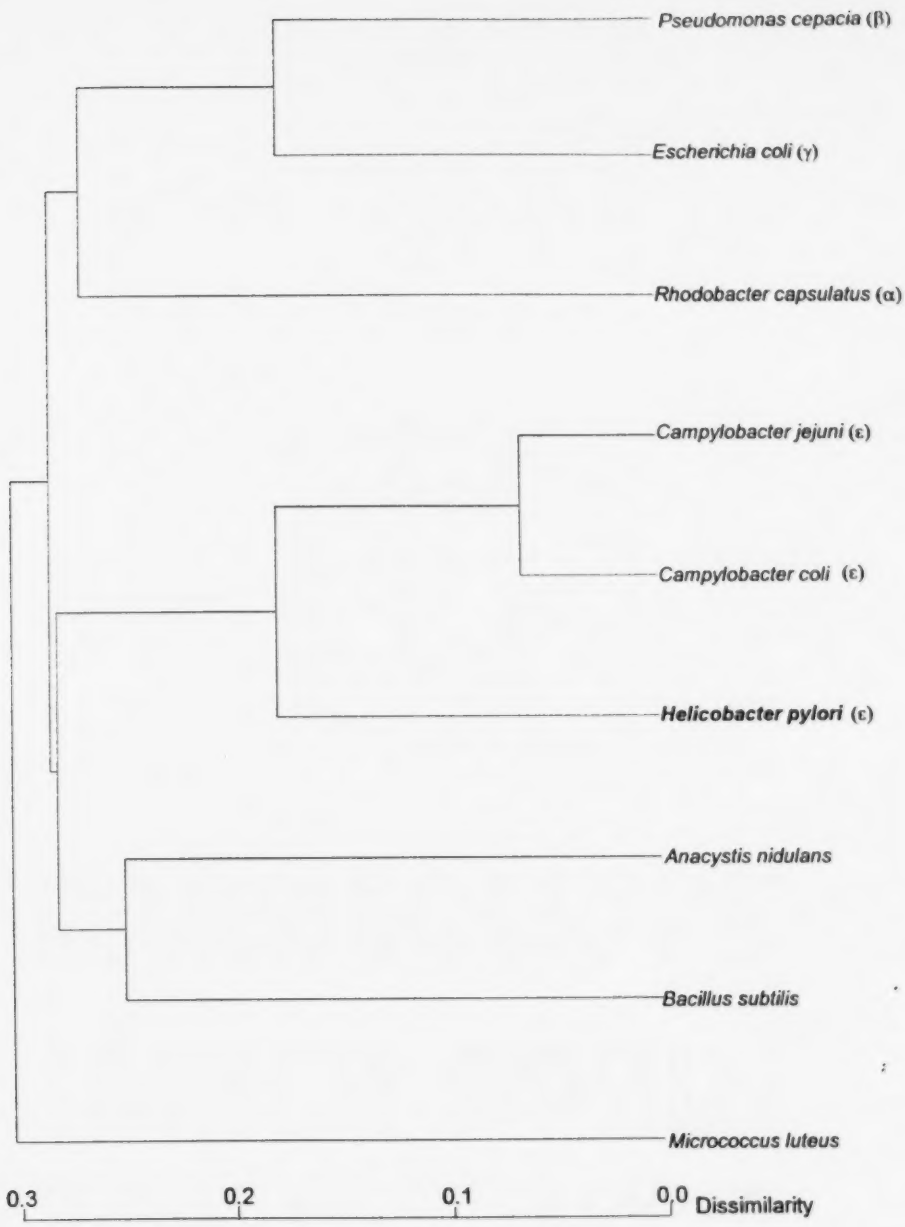


Figure 4-14. Phylogenetic tree based on comparisons of complete 23S rRNA gene sequences, including representative organisms from four of the five *Proteobacteria* subdivisions and representative organisms from the Gram-positive (*B. subtilis* and *M. luteus*) and cyanobacteria (*A. nidulans*) phyla

Proteobacteria subdivisions are shown in brackets.

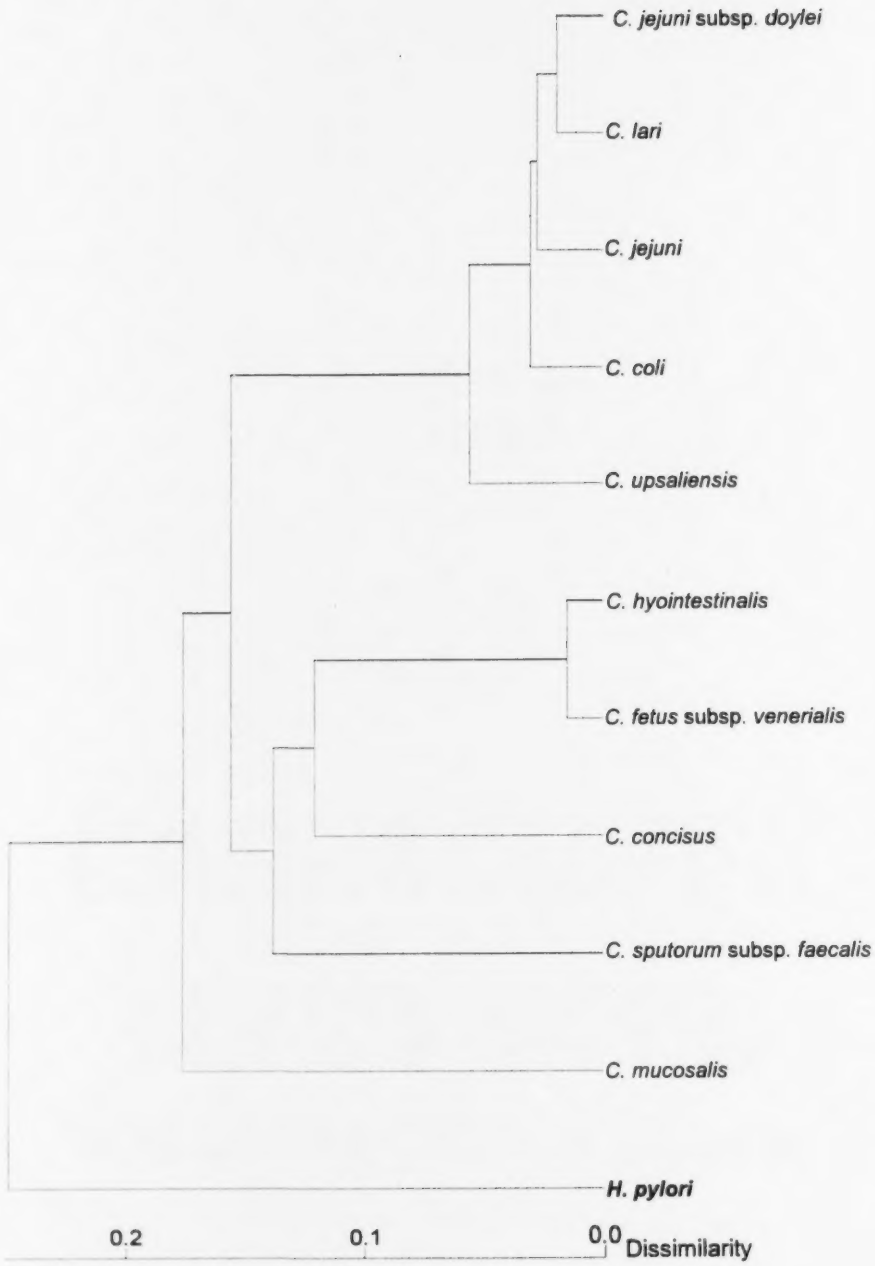


Figure 4-15. Phylogenetic tree based on comparisons of available partial 23S rRNA gene sequences of *Campylobacter* species and *H. pylori*

(Trust, *et al* ., 1994) 23S genes homologous to the available *Campylobacter* partial sequences were used.

The pileup program of GCG was also used to construct the dendogram comparison of the complete 5S gene sequence of *H. pylori* UA802 with the complete 5S gene sequences from representative organisms of all 5 subdivisions of the class *Proteobacteria*, as well as eight out of ten of the other Eubacterial classes described by Woese (1987). Figure 4-16 shows the dendogram constructed for this comparison.

D. Discussion

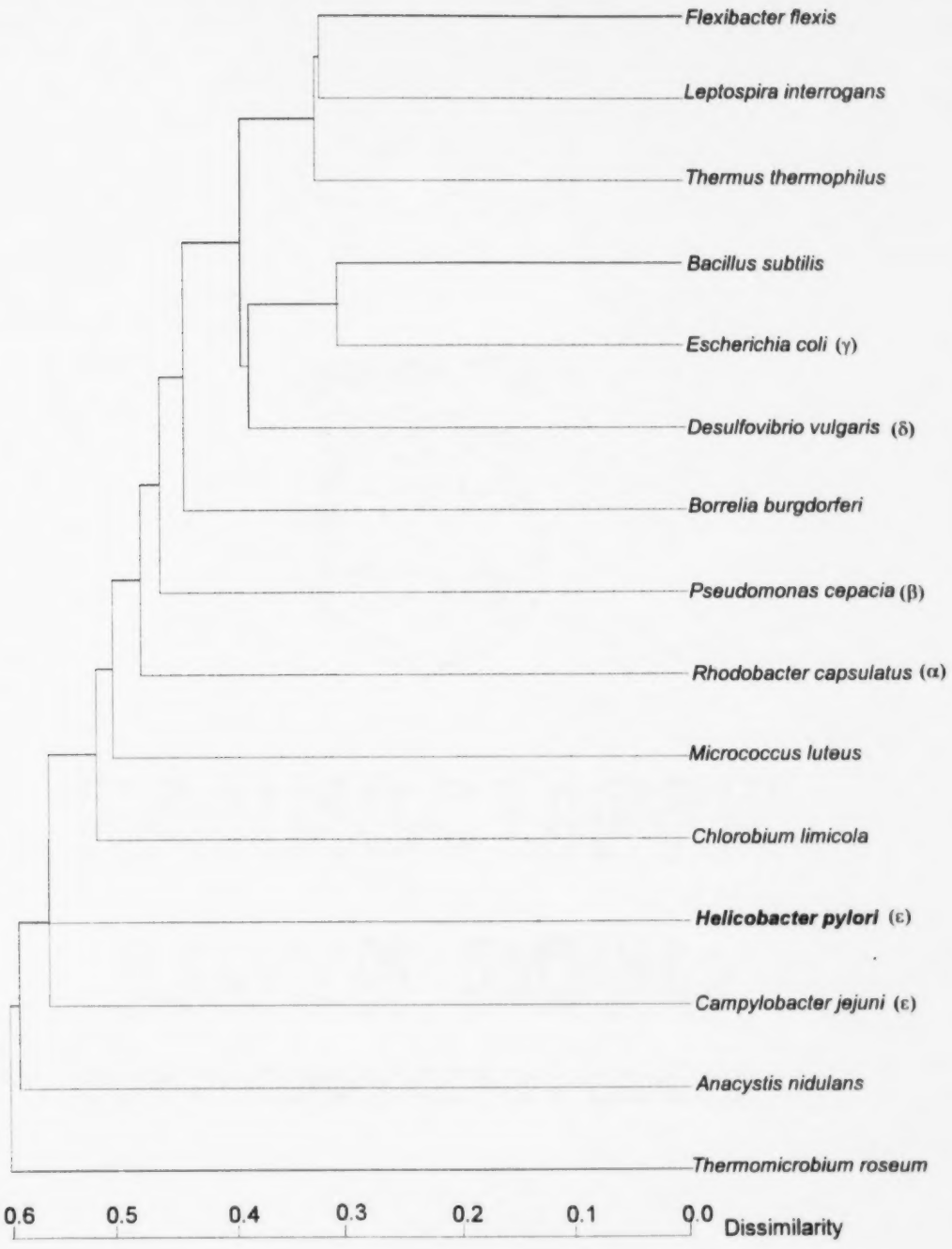
D.1. Evidence for a P-type ATPase in *H. pylori* UA802 and UA763

The existence of a P-type ATPase in *H. pylori* was shown by the hybridization of chromosomal digestions of two strains of *H. pylori*, UA802 an UA763, with cloned P-type ATPases from the *E. hirae* (K^+-ATPase), *E. coli* (*kdp*b) and rat ((H^+-K^+)-ATPase gene)(Figure 4-3). This suggested that *H. pylori* encoded an element homologous to P-type ATPases. Hybridization was very weak, which was not surprising as the homology at the DNA level between these ATPases was low, even in regions which showed extensive amino acid conservation (Figure 4-1). The sizes of the hybridizing bands in the two strains differed. Other genes in *H. pylori* have shown differences in restiriction patterns. Both urease genes (*ure*ABCD) and flagellin genes (*fla*A) have shown restriction fragment length polymorphism (RFLP) by PCR (Foxall *et al.*,

Figure 4-16. Phylogenetic tree based on comparisons of complete 5S rRNA gene sequences from representative organisms of eight of the ten Eubacterial phyla

Proteobacteria subdivisions are indicated in brackets.

Sequences used in comparisons:
Flexibacter flexis 5S- Van den Eynde *et al.*, 1990.
Accession #M33889
Leptospira interrogans 5S- Fukunaga *et al.*, 1990.
Accession #D90074
Thermus thermophilus 5S- Komiya *et al.*, 1983.
Accession #X01554
Bacillus subtilis 5S- Marotta *et al.*, 1976.
Accession #M10815
Escherichia coli 5S- Brownlee *et al.*, 1967.
Accession #K00609
Desulfovibrio vulgaris 5S- Miura *et al.*, 1986.
Accession #M32519
Borrelia burgdorferi 5S- Gazumyan *et al.*, 1994.
Accession #U03396
Pseudomonas cepacia 5S- Vandenberghe *et al.*, 1985.
Accession #X02629
Rhodobacter capsulatus 5S- Kato and Komagata, 1986.
Accession #X03904
Micrococcus luteus 5S- Hori *et al.*, 1980.
Accession #J10868
Chlorobium limicola 5S- Van den Eynde *et al.*, 1990.
Accession #M33893
H. pylori UA802 5S- this work; Figure 4-10, bases 3569-3711
Accession # U27270
Campylobacter jejuni 5S- Kim *et al.*, 1994.
Accession #Z29326
Anacystis nidulans 5S- Corry *et al.*, 1974.
Accession #X00757
Thermomicrobium roseum 5S- Van den Eynde *et al.*, 1990.
Accession #M33891


Flexibacter flexis
Leptospira interrogans
Thermus thermophilus
Bacillus subtilis
Escherichia coli (γ)
Desulfovibrio vulgaris (δ)
Borrelia burgdorferi
Pseudomonas cepacia (β)
Rhodobacter capsulatus (α)
Micrococcus luteus
Chlorobium limicola
Helicobacter pylori (ε)
Campylobacter jejuni (ε)
Anacystis nidulans
Thermomicrobium roseum
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Dissimilarity

1992; Akopyants *et al.*, 1995). In addition, Cover *et al* (1994) showed that there could be variation in the *vac*A sequences from stains which produced active forms of the VacA protein and those which produced inactive forms. The hybridization data presented here using the P-type ATPase probes suggest that similar variations may occur in these genes between strains of *H. pylori*.

Cloning of a P-type ATPase from *H. pylori* UA802

Amino acid comparisons have shown that certain regions of P-type ATPases are highly conserved (Hesse *et al.*, 1984; Solioz *et al.*, 1987; Shull and Lingrel, 1986). PCR primers were designed from two of these conserved regions (Figure 4-2). The primers were designed from the nucleotide sequences shown. At positions where the nuclotide sequences were variable, the *E. hirae* sequence was used, as this organism has similar a G +C ratio as *H. pylori*. A 600 bp P-type ATPase PCR probe was amplified from *H. pylori* UA802 using these primers and was used to screen the *Bgl*II and *Sau*3AI λ-DASH II genomic libraries of *H. pylori* UA802. Twenty-three clones were identified with this probe and two of these clones, λP9 and λP14, were further characterized by restriction endonuclease digestion analysis and hybridization. The PCR probe hybridized to a 1.5 kb *Eco*RI-*Hind*III fragment from λP9 and a 2.7 kb *Eco*RI-*Hind*III fragment from clone λP14. The 2.7 kb *Eco*RI-*Hind*III fragment from clone λP14 was subcloned and used in the characterization of the putative P-type ATPase by Ge *et al.*, 1995. Ge *et al.*

found that the cloned *H. pylori* UA802 DNA encoded 3 open reading frames (ORF), designated ORF1, ORF2, and ORF3. The polypeptide encoded by ORF1 was found to have extensive homology with several bacterial and eucaryotic P-type ATPases, most strikingly with ATPases involved in copper transport (Vulpe *et al.*, 1993; Odermatt *et al.*, 1993), and was designated *hpCop*A. The polypeptide encoded by ORF2 was found to have homology with the mercury binding protein merP (Barrineau *et al.*, 1984) and was designated *hpCop*P. No homology could be determined for the ORF3 polypeptide. A knock-out mutant of the *hpCop*A gene was constructed by insertion of a chloramphenicol-resistance cassette, disrupting the gene. This knock-out mutant was susceptible to omeprazole and a number of metal ions at levels similar to those of the wild-type, but was shown to have increased sensitivity to Cu^{2+} ions. Based on the experimental results and the homology of *hpCop*A to a copper transport gene, it was concluded by Ge *et al.* (1995) that this P-type ATPase is probably involved in copper transport in *H. pylori* UA802.

D.2. Cloning and characterization of the rRNA genes of *H. pylori* UA802

When the genomic libraries of *H. pylori* UA802 were screened for the 16S and 23S genes with PCR probes, 18% and 7.5% of the plaques hybridized with the 16S and 23S probes respectively. This was significantly higher than the expected frequency of approximately 2% calculated based on an average insert fragment size of 17 kb, a genome size of 1700

kb and the existence of 2 copies of the rRNA genes. This suggested that clones carrying the rRNA genes or regions flanking the rRNA genes had some selective advantage over other clones. Similar results were obtained in the construction of genomic libraries of the nematode *Caenorhabditis elegans*, where ribosomal DNA (rDNA) sequences were cloned at a frequency higher than expected using both lambda vectors (Ellis *et al.*, 1986) and cosmid vectors (Gibson *et al.*, 1987). Gibson *et al.* suggested that the rDNA had greater stability when cloned in *E. coli*, as it does not encode a protein and thus is less likely to be toxic. In addition, Gibson *et al.* found that the G + C content of the rDNA (49.3 mol%) was much higher than the rest of the genome (36 mol%), which may have reduced sequence-related instability of the cloned DNA. The cloned 23S-5S regions of *H. pylori* UA802 also showed a higher G + C ratio (45.3 mol%) than the ratio calculated for the genome (35.2 mol%; Béji *et al.*, 1988). As a result, clones harboring the rRNA genes may have been more stable than other clones. Since the plates that were screened were not primary plates, but were plates made using pooled overlays from primary plates, this phenomenon may have been amplified. Although it was not tested, the frequency of rRNA clones on the primary plates may have been closer to the expected number.

Characterization of the rRNA genes of *H. pylori* UA802

The 16S rRNA sequences of many diverse organisms have been extensively studied and the phylogenetic relationships

between many organisms have been elucidated by 16S rRNA sequence comparisons (Woese, 1987). In contrast, there have been relatively few studies comparing 23S. As of 1992, only twenty-one complete eubacterial 23S sequences had been reported (Gutell *et al.*, 1992) as opposed to 625 complete eubacterial 16S sequences (de Rijk *et al.*, 1992). Although quite a few 5S sequences have been reported (Erdmann *et al.*, 1985), the small size of the 5S molecules makes them less suitable for phylogenetic comparisons (Woese, 1987). Since the partial and complete sequences of the 16S rRNA genes of a number of strains of *H. pylori* have already been determined (Romaine *et al.*, 1987; Eckloff *et al.*, 1994) my work was focussed on the characterization of the 23S and 5S rRNA genes of *H. pylori* UA802.

The complete nucleotide sequences of both copies of the 23S and 5S genes of *H. pylori* UA802 were determined and are shown in Figure 4-10. The fact that the sequences of both copies of the 23S and 5S genes, as well as the flanking and intervening regions, from *H. pylori* UA802 were identical is not unusual. Comparisons of the sequences of 23S and 5S genes from six different *rrn* operons from *Bacillus subtilis* showed very little variation in the 23S and 5S sequences and the intervening sequences (Appendix C). All six copies of the 23S and 5S genes were nearly identical with minor variations in 2 copies (*rrnJ* and *rrnW*). The 5'-flanking regions to the 23S genes for the *rrnO* and *rrnB* were identical at for at least 327 bp, as were the 5' flanking regions of

*rrn*A and *rrn*J (Appendix C). The regions 3' to the 5S gene showed more variation, but the *rrn*I and *rrn*O sequences were identical up to 39 bp past the end of the 5S gene (Appendix C). Similarly, the regions 5' to the 23S genes of *H. pylori* UA802 were identical for at least 372 bases before the 23S gene and, by comparison of patterns on the sequencing gels (not shown), were similar several hundred bases beyond the point to which sequences were determined. The regions 3' to the 5S genes were identical up to 61 bp from the end of the 5S genes, at which point the sequences diverged. Although the sequences of individual copies of the 16S rRNA genes of *H. pylori* have not been compared, sequencing data obtained by Eckloff *et al.* (1994) would suggest that very little variation occurs between individual copies of the 16S genes in *H. pylori*. In this study the DNA sequence of the 16S genes of five different strains of *H. pylori* was determined. The 16S genes were first amplified from the chromosomes by PCR and the sequence was determined from these PCR products. If any variation was present in the individual copies of the 16S gene, this would become apparent, as ambiguous sequence would be obtained. No such ambiguities were observed for any of the 16S sequences obtained. In addition, although some variations were observed between the 16S genes from different strains of *H. pylori*, these variations were small (0.2-0.5%). As Eckloff *et al.* state, this would indicate that the genomic variability observed between strains of *H. pylori* does not extend to the 16S genes. The data presented here for the 23S

and 5S genes from *H. pylori* UA802 also suggest that variability is not likely to occur in these regions.

Generally, the rRNA genes of procaryotes are found clustered together in operons. In some organisms, as mentioned in Chapter 1, the rRNA genes are not linked. Some strains of *Borrelia* spp.(Schwartz *et al.*, 1992), *C. jejuni* and *C. coli* (Taylor *et al.*, 1992b), *Mycoplasma hyopneumoniae* (Taschke *et al.*, 1986), *Mycoplasma gallisepticum* (Chen and Finch, 1989), *Leptospira interrogans* (Fukunaga and Mifuchi, 1989), *Pirellula marina* (Liesack and Stackebrandt, 1989), *Thermoplasma acidophilum* (Ree and Zimmermann, 1990), *Thermus thermophilus* (Hartman and Erdmann, 1989) and *H. pylori* (Taylor *et al.*, 1992a; Bukanov and Berg, 1993) are examples of procaryotes in which the rRNA genes are not linked.

Sequence of the 5' region of the 23S genes of *H. pylori* UA802 did not reveal the presence of a 16S gene (Figure 4-10). Thus the rRNA genes of *H. pylori* UA802 do not appear to be linked. This was consistent with results obtained for the construction of the genomic restriction map of *H. pylori* UA802 (Taylor *et al.*, 1992a; Chapter 5, this thesis) where the 16S rRNA genes were not linked with the 23S genes. Coordination of the amounts of rRNA species in organisms where the rRNA genes are linked appears to be governed by the co-transcription of the genes in operons, followed by post-transcriptional processing to give the individual species of rRNA (Nomura *et al.*, 1984). In *E. coli*, transcription of the rRNA operons is regulated by two tandem promoters (Nomura *et*

al., 1984). These promoters are found approximately 200-300 bases upstream from the 16S rRNA gene and are normally found approximately 120 bp apart (Gilbert *et al.*, 1979). Other regions surrounding the promoters, such as the upstream activation region (UAR) and the factor-dependent and factor-independent regions, are also important in regulating the transcription of rRNA in *E. coli* (Condon *et al.*, 1992).

Three Pribnow boxes (consensus sequence is TATAAT) could be found 5' to the start of the *H. pylori* UA802 23S gene (Figure 4-10). No −35 region (consensus sequence is GTTGAC) could be clearly defined 5' to any of these Pribnow boxes. No sequences homologous to the UAR, factor-dependent or factor-independent regions could be found. Thus, the control mechanisms for *H. pylori* UA802 rRNA genes must differ from those found in *E. coli* and would require further investigation. It is also possible that sequences similar to the UAR, factor-dependent or factor-independent regions existed beyond the regions which were sequenced for this research. Cloning and sequencing the 5' regions of the 23S genes and *in vitro* transcription would provide more information about the transcriptional control of this operon. It is interesting to note that the Pribnow box designated P1 was exactly ten bases from the start of the mature 23S sequence, which was determined by primer extension (Figure 4-11). In *E. coli*, transcription begins almost 300 bp upstream from the 16S gene (de Boer *et al.*, 1979). This suggests that, unlike *E. coli*, the 5'-end of the 23S-5S transcript is

not processed and transcription, at least from the P1 promoter, begins at position 373 (Figure 4-10), which correlates with the primer extension data (Figure 4-11).

The end of the 23S gene and the start of the 5S gene are separated by 229-235 bp (Figure 4-10) and no promoter sequences could be identified in this region. The 23S and 5S genes must therefore be co-transcribed and post-transcriptionally modified to give the mature 23S and 5S rRNA molecules. It appears that the processing mechanism is not highly specific, as a total of 5 primer extension products were observed for the 5'-end of the 5S rRNA molecule (Figure 4-11). Putative end sites were determined by sequence alignments of the 23S and 5S sequences of a number of other procaryotes with the sequences obtained here (data not shown).

Phylogenetic comparisons of 23S and 5S sequences

As mentioned in Chapter 1, *H. pylori* has been placed in the epsilon subdivision of the class *Proteobacteria* along with *Campylobacter* spp., *Arcobacter* spp., *Wolinella* spp., and *Flexispira* spp. (Vandamme *et al.*, 1991). Phylogenetic comparisons based on 23S sequence are limited by the fact that, compared to 16S rRNA sequences, relatively few complete 23S sequences have been determined. Recently the complete sequences of the 23S genes of *C. jejuni* (Kim *et al.*, 1993, GenBank accession number Z29326) and *C. coli* (Trust *et al.*,

1994; GenBank accession number U09611) have been determined. Partial sequences of the 23S genes of a number of species of *Campylobacter* are also available (Van Camp *et al.*, 1993).

The results of the comparisons of the *H. pylori* UA802 23S gene with the complete sequences of the 23S genes of *C. jejuni*, *C. coli* and representative organisms from the other subdivisions of the phylum *Proteobacteria* (Figure 4-12) were consistent with earlier phylogenetic classifications of *H. pylori*. The comparisons showed that *H. pylori*, along with *C. jejuni* and *C. coli*, belongs in a separate subdivision (Trust *et al.*, 1994). Gram-positive bacteria and cyanobacteria classes are most closely related to the *Proteobacteria* (Woese, 1987). Therefore, *Bacillus subtilis* (Gram-positive, low G + C), *Micrococcus luteus* (Gram-positive, high G + C) and *Anacystis nidulans* (cyanobacteria) were included in the phylogenetic tree (Figure 4-12).

The results of the comparison of the partial 23S sequence of *H. pylori* UA802 with the partial 23S sequences of several *Campylobacter* species determined by van Campe *et al.* (1993) were consistent with a previous study which showed that *H. pylori* was distinct from the *Campylobacter* spp. (Vandamme et al., 1991). In this study, phylogenetic relationships were based on melting temperatures and percent rRNA binding in DNA-rRNA hybrids. The phylogenetic positioning of *H. pylori* was based on DNA-rRNA hybrids using labelled-rRNA from *C. fetus* subsp. *fetus*, *C. concisus*, *C. coli*, and *Wolinella succinogenes*. Reciprocal hybrids using rRNA from

H. pylori were not performed. The phylogenetic comparisons done for this research provide direct evidence of the relationships *H. pylori* has with several members of the genus *Campylobacter* and supports the findings of Vandamme *et al.* (1991) that *H. pylori* belongs in a separate phylogenetic group from the *Campylobacter* spp..

Comparisons of 5S gene sequences showed that *H. pylori* was phylogenetically distinct from the other members of the class *Proteobacteria* (Figure 4-14). Although a close association between *H. pylori* and *C. jejuni* could not be established, *H. pylori* and *C. jejuni* were shown to be equally dissimilar to other organisms based on the 5S comparisons. Complete 5S gene sequences from organisms of eight of the ten phyla, described by Woese (1987) are also represented on this phylogenetic tree, including representative organisms from all 5 of the subdivisions of the phylum *Proteobacteria*. The relationships between the *Proteobacteria*, Gram-positives (*B. subtilis*, *M. luteus*), and cyanobacteria (*A. nidulans*) observed based on the 23S rRNA sequences could not be demonstrated with the phylogenetic comparisons based on the 5S rRNA sequences. This reflects the lower resolving power of comparisons using the 5S gene sequences as compared to the use of 16S or 23S gene sequences. Because of their relatively small size compared to the 16S and 23S rRNA, differences in 5S rRNA sequences would be more useful for resolution of more recent evolutionary change (Woese, 1987). Similarly, since 23S rRNA molecules are much larger than the

16S rRNA and have approximately twice the number of domains (stem-loop structures), they provide more accurate phylogenetic comparisons than the 16S rRNA (Woese, 1987).

Chapter 5

Pulsed field gel electrophoresis analysis and revision of the genomic map of *H. pylori* strain UA802

A. Introduction

A genomic restriction map of *H. pylori* UA802 (Figure 5-1) was constructed in 1992 using pulsed-field gel electrophoresis with the enzymes *Not*I and *Nru*I (PFGE, Taylor *et al.*, 1992a). PFGE analyses done for this research led to the revision of this map and the addition of *Sfi*I sites.

Genomic restriction maps, in general, are useful for studies of genomic variability between strains of bacteria, as well as for the localization of genes and mutant phenotypes (Weinstock, 1994). The genomes of different *H. pylori* strains have been shown to be highly variable (Majewski *et al.*, 1988; Clayton *et al.*, 1989; Akopyantz *et al.*, 1992; Taylor *et al.*, 1992a). Therefore, genomic restriction maps of *H. pylori*, such as the map constructed for this work, can be used for comparisons of genetic organization of the chromosomes of different strains. Genomic restriction maps of a number of strains of *H. pylori* have been constructed (Bukanov and Berg, 1994; Jiang, 1994). Alignment of these maps with the map of *H. pylori* UA802 constructed for this research may give insights about the genomic variability observed between strains of *H. pylori*.

H. pylori UA 802

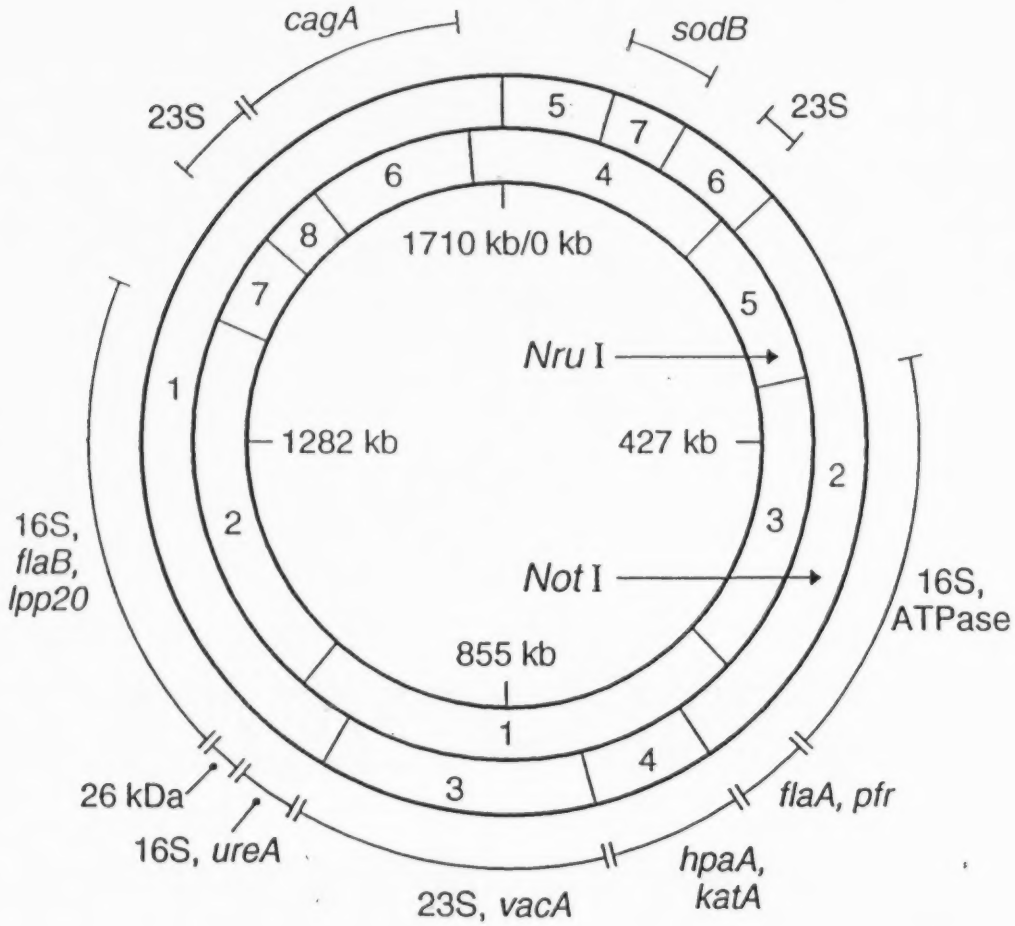


Figure 5-1. Restriction map of *H. pylori* UA802 constructed by Taylor *et al.*, 1992a

B. Materials and Methods

Please refer to Chapter 2 for detailed methodology used for this chapter.

C. Results

C.1. Construction of the *H. pylori* UA802 restriction map

PFGE analyses

PFGE analysis of *H. pylori* UA802 chromosomal DNA was the initial step in the construction of the genomic restriction map. Embedded *H. pylori* UA802 DNA was restricted with *Not*I, *Nru*I, and *Sfi*I as described in Chapter 2 (Section D). PFGE of the digested DNA was carried out with both regular agarose and LMP agarose. Complete digestion of the *H. pylori* UA802 DNA with these enzymes gave nine *Nru*I fragments, eight *Not*I fragments and three *Sfi*I fragments. The sizes of the various restriction fragments were estimated from the PFGE analyses of these complete digestions (Table 5-1). The size of the *H. pylori* UA802 genome was estimated by totaling the sizes of the fragments for each enzyme. The average size estimated from these totals was 1695 ± 17 kb and was in agreement with the size estimated previously for *H. pylori* UA802 (Taylor *et al.*, 1990) and with estimated sizes of other strains of *H. pylori* (Bukanov and Berg, 1994; Jiang, 1994), which ranged from 1674 kb to 1730 kb.

Genomic restriction map construction

To determine the order of the fragments in the genomic

Table 5-1. Restriction fragment sizes of *H. pylori* UA802 DNA

Frag. #	*Not*I fragments	*Nru*I fragments	*Sfi*I fragments
1	710.0 kb ± 15.0	420.0 kb ± 6.0	870.0 kb ± 15.0
2	458.0 kb ± 8.0	351.0 kb ± 5.0	458.0 kb ± 7.0
3	200.0 kb ± 4.0	253.0 kb ± 8.0	340.0 kb ± 9.0
4	89.0 kb ± 4.0	222.0 kb ± 3.0	-
5	75.0 kb ± 2.0[a]	152.0 kb ± 6.0	-
6	75.0 kb ± 2.0[a]	123.0 kb ± 8.0	-
7	54.0 kb ± 1.0	87.0 kb ± 2.0	-
8	50.0 kb ± 3.0	55.0 kb ± 3.0	-
9	-	38.0 kb ± 2.0	-
Total	1711.0 kb	1704.0 kb	1668.0
Average size	**1695.0 kb ± 17**		

a- fragments 5 and 6 formed a doublet band

restriction map, Southern blots of PFGE gels of complete digestions of *H. pylori* DNA were hybridized with restriction digestion fragments isolated from LMP agarose. The results of the fragment hybridizations are shown in Table 5-2. The hybridization patterns of some of the fragments differed from the data obtained in earlier map construction of *H. pylori* UA802 (Taylor *et al.*, 1992). These differences are as follows and are also indicated on Table 5-2. *Nru*I fragment 1 hybridized with *Not*I fragments 1, 2, 4, and 6, while on the earlier map *Nru*I fragment 1 hybridized with *Not*I fragments 1, 2, 3, and 4. *Nru*I fragment 3 hybridized with *Not* I fragments 2 and 5, while on the earlier map *Nru*I fragment 3 hybridized only with *Not*I fragment 2. *Nru*I fragment 4 hybridized with *Not*I fragments 1, 3, 5, and 8 (formerly 7), while on the earlier map *Nru*I fragment 4 hybridized with *Not*I fragments 1, 5, 6 (7 on revised map), and 7 (8 on the revised map). *Nru*I fragment 5 hybridized with *Not*I fragment 3, while on the earlier map *Nru*I fragment 5 hybridized with *Not*I fragments 2 and 6 (7 on the revised map). *Nru*I fragment 6 hybridized with *Not*I fragment 2, while on the earlier map *Nru*I fragment 6 hybridized with *Not*I fragment 1. *Nru*I fragment 8 hybridized with *Not*I fragments 3 and 6, while on the earlier map *Nru*I fragment 1 hybridized with *Not*I fragment 1. *Not*I fragment 1hybridized with *Nru*I fragments 1, 2, 4, and 7, while on the earlier map *Not*I fragment 1 hybridized with *Not*I fragments 1, 2, 4, 6, 7, and 8. *Not*I fragment 2 hybridized

Table 5-2. *H. pylori* UA802 fragment localization

Digest Fragment	*NruI**	*NotI**	*SfiI***
NruI 1	1	1,2,4,7 (1,2,3,4)	1
NruI 2	2	1	1,2
NruI 3	3	2,5 (2)	1
NruI 4	4	1,3,5,8 (1,5,6,7)	2,3
NruI 5	5	3 (2,6)	1,3
NruI 6	6	2 (1)	1
NruI 7	7	1	2
NruI 8	8	3,5 (1)	1
NruI 9**	9	3	3
NotI 1	1,2,4,7 (1,2,4,6,7,8)	1	1,2,3
NotI 2	1,3,6 (1,3,5)	2	1
NotI 3	4,5,8,9 (1)	3	1,3
NotI 4	1	4	1
NotI 5	4	5	3
NotI 6	3,8	5	1
NotI 7	1 (4,5)	6	1
NotI 8	4	7	3
SfiI 1**	1,2,3,5,6,8	1,2,3,4,6,7	1
SfiI 2**	2,4,7	1	2
SfiI 3**	4,5,9	1,3,5,8	3

* Numbers in brackets indicate fragments which hybridized in the original map (Taylor *et al.*, 1992)

** Did not appear on original map

with *Nru*I fragments 1, 3, and 6, while on the earlier map *Not*I fragment 2 hybridized with *Not*I fragments 1, 3, and 5. *Not*I fragment 3 hybridized with *Nru*I fragments 4, 5, 8, and 9, while on the earlier map *Not*I fragment 3 hybridized with *Not*I fragment 1. *Not*I fragment 7 (6 on the earlier map) hybridized with *Nru*I fragment 1, while on the earlier map *Not*I fragment 7 hybridized with *Not*I fragments 4 and 5. In addition, *Nru*I fragment 9 was not included in the original map and *Not*I fragment 5 from the original map was shown to be a doublet. Figure 5-2 shows hybridization data confirming the results presented in Table 5-2. The data from the fragment hybridizations were compiled to construct the revised restriction map of *H. pylori* UA802 (Figure 5-3).

On the revised map, the relative positions of *Not*I fragments 3 and 7 (previously 6) are different from the original map. As well, the relative positions of *Nru*I fragments 6 and 8 are different from the original map.

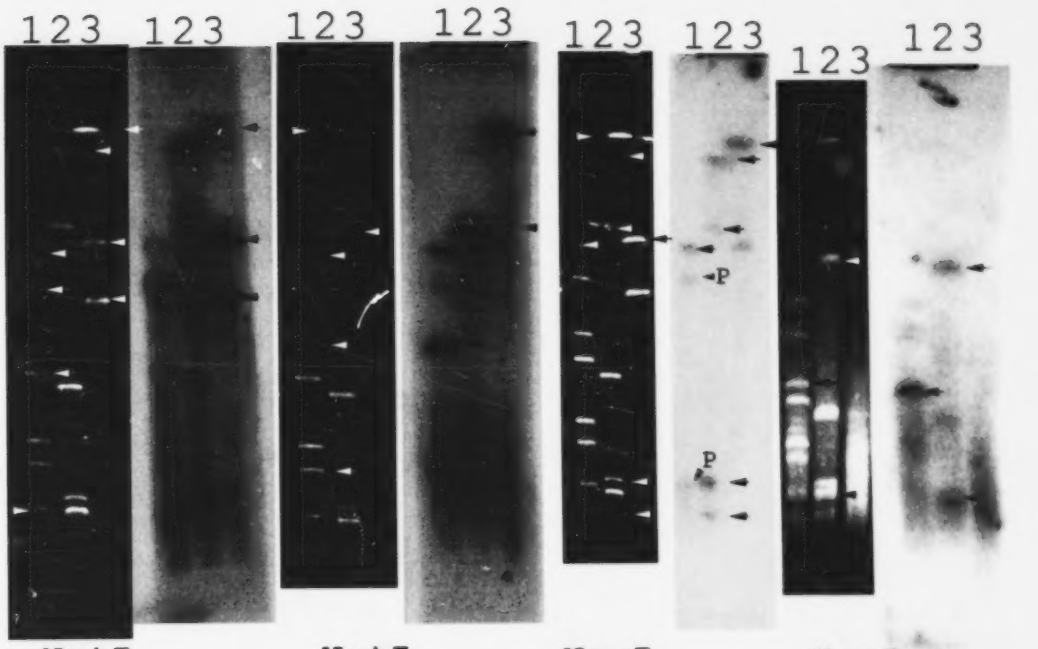
Localization of genes on the *H. pylori* UA802 map

The Southern blots were also probed with a variety of genetic elements to localize them on the map. These probes and their sources are listed in Table 5-3. Hybridization data confirmed that *H. pylori* UA802 has two copies of the 16S and 23S rRNA genes. This is in agreement with the findings of Bukanov and Berg (1994) and Jiang (1994), who found that other strains of *H. pylori* had two copies of each rRNA gene. Hybridization data combined with sequence data (Chapter 4) showed that the two 23S genes straddled the *Nru*I sites

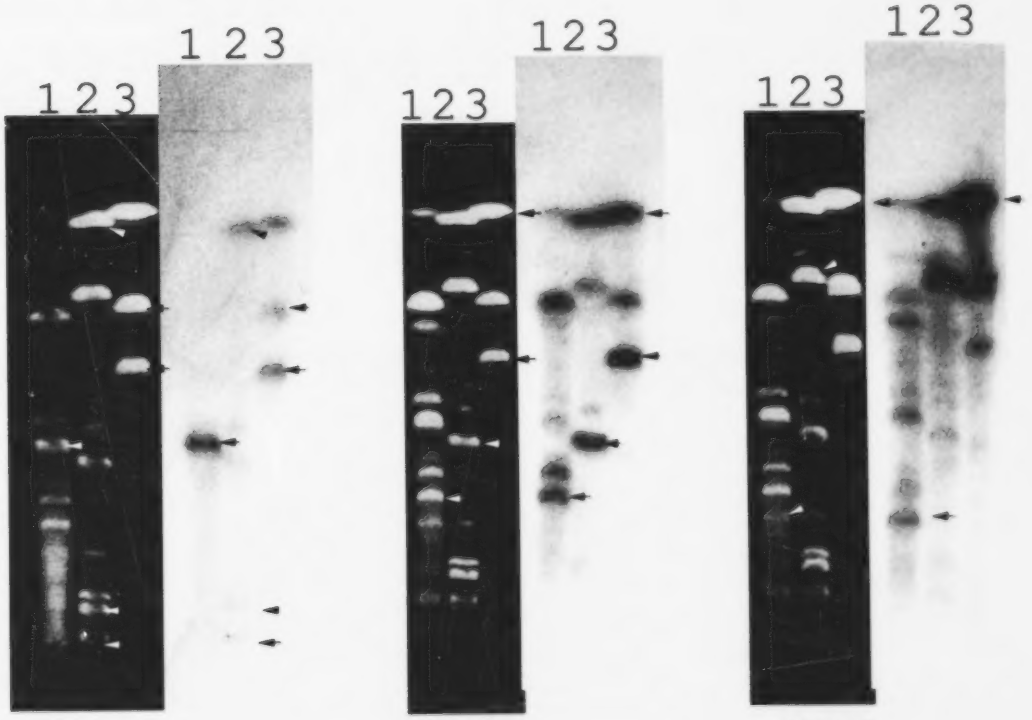
Figure 5-2. PFGE of *H. pylori* UA802 DNA digested with *Not* I, *Nru*I, and *Sfi*I, and Southern blots of the gels probed with α-[^{32}P]-dATP-labelled PFGE restriction digestion fragments isolated using LMP-agarose

1- *Nru*I digest of *H. pylori* UA802
2- *Not*I digest of *H. pylori* UA802
3- *Sfi*I digest of *H. pylori* UA802

Each panel represents an agarose gel on the left and the Southern transfer of that gel, probed with the PFGE fragment indicated, on the right. Arrows on the agarose gel show fragments which hybridized. Fragments on the autoradiograms which are not indicated with an arrow were presumed to be partial digest fragments. For the *Nru*I fragment 1 hybridization, the probe itself contained a partial digest fragment, accounting for the hybridization of *Nru*I fragments 2 and 7 (indicated with a P on the Southern blot). PFGE was carried out in 0.75% agarose at 8°C with a constant voltage of 190v, with pulse times of 8h at 10s, 8h at 30s and 8h at 60s for a total running time of 24h. PFGE fragments used as probes were run under the same conditions through 1% LMP-agarose, visualized by ethidium bromide staining, excised from the gels and labelled with α-[^{32}P]-dATP by nick translation. Hybridizations were carried out at 37°C in the presence of 50% formamide.



NotI Fragment 1 | NotI Fragment 2 | NruI Fragment 1 | NruI Fragment 3



NruI Fragment 4 | NruI Fragment 5 | NruI Fragment 6

H. pylori UA 802

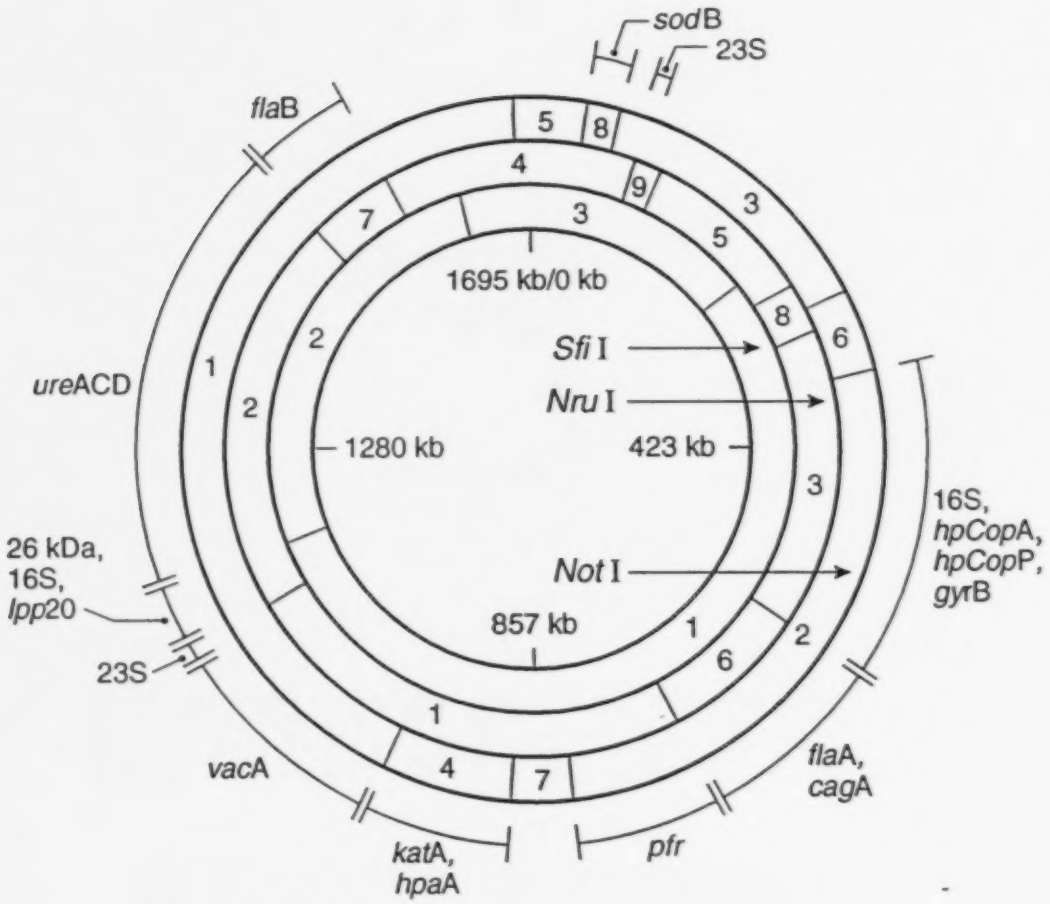


Figure 5-3. Revised genomic restriction map of *H. pylori* UA802

Table 5-3. Probes used for localization of genes in the *H. pylori* UA802 genome

Gene	Plasmid or PCR probe	Origin	Source
23S rRNA	PCR	*H. pylori* UA802	this work
16S rRNA	PCR	*H. pylori* UA802	this work
*ure*ACD	pILL594	*H. pylori* 85P	Labigne *et al.*, 1991
26 kDa protein	p26K	*H. pylori* 915	O'Toole *et al.*, 1991
*vac*A	pCTB1 PCR	*H. pylori* 60190 *H. pylori* UA802	Cover *et al.*, 1994 this work
prf	PCR	*H. pylori* UA802	Jiang, 1994
*fla*A	PCR	*H. pylori* UA802	Jiang, 1994
*fla*B	PCR	*H. pylori* UA802	Jiang, 1994
*hpCop*A	PCR	*H. pylori* UA802	Jiang, 1994; this work
*sod*B	pCS188	*H. pylori*	Spiegelhalder *et al.*, 1993
*cag*A	PCR	*H. pylori* UA802	Jiang, 1994
*lpp*20	PCR	*H. pylori* UA802	Jiang, 1994
*gyr*A	plasmid	*H. pylori*	Moore *et al.*, 1995
*gyr*B	PCR	*H. pylori* UA802	Jiang, 1994
*kat*A	PCR	*H. pylori* UA802	Jiang, 1994
*hpa*A	PCR	*H. pylori* UA802	Jiang, 1994

For probes which have the source indicated as "this work" and "Jiang, 1994", the PCR probes were designed from published sequences of the genes. See Appendix B for the primers used.

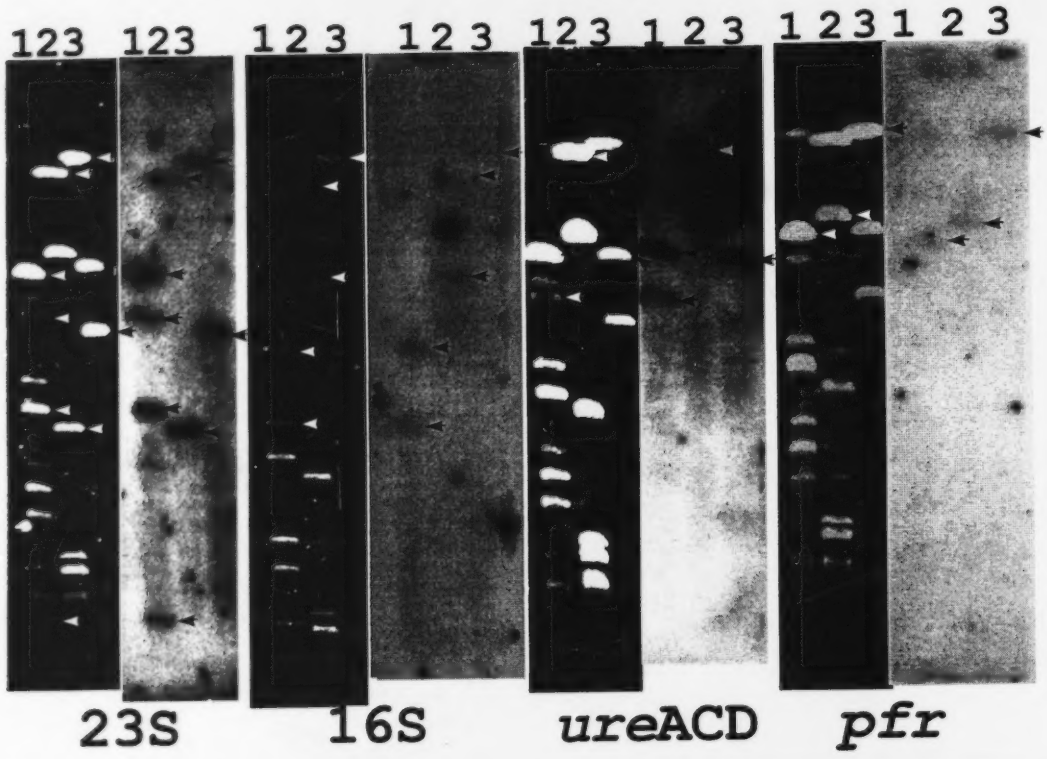
between *Nru*I fragments 1 and 2 and *Nru*I fragments 4 and 9 (Figure 5-3). During the construction of the original map, three other genes (*ure*ACD, 16S and 26 kDa-protein genes) all appeared to straddle the site between *Nru*I fragments 1 and 2. It was found that *Nru*I fragments 2 and 7 often form a partially digested fragment, particularly resistant to complete digestion, which is the same size as *Nru*I fragment 1. Hybridization data later showed that the *ure*ACD genes were localized to fragment *Nru*I 2 and that the 16S and 26 kDa-protein genes were localized to *Nru*I fragment 1 (Figure 5-4, 26 kDa-protein data not shown). The *gyr*A gene could not be localized for undetermined reasons.

Assignments of map positions of the various genetic elements, shown in Figure 5-3, were based on hybridization data presented in Table 5-4. The positions of some of the genetic elements on the revised map differed from their positions on the previous map. As mentioned earlier, the positions of the 23S, 16S, *ure*ACD and the 26 kDa protein genes were revised. The positions of the *vac*A, *fla*A, and *fla*B genes were also different. The *vac*A gene was localized to *Nru*I fragment 1 and *Not*I fragment 3 on the original map, but localized to *Nru*I fragment 1 and *Not*I fragment 1 on the revised map. The *fla*A gene originally was localized to *Nru*I fragment 1 and *Not*I fragment 2, but on the revised map was found to hybridize to *Nru*I fragment 6 and *Not*I fragment 2. The *fla*B gene was mapped to *Nru*I fragment 2 and *Not*I fragment 1 on the original map, but was later mapped to *Nru*I

Figure 5-4. PFGE of *H. pylori* UA802 DNA digested with *Not* I, *Nru*I, and *Sfi*I, and Southern blots of the gels hybridized with various α-[^{32}P]-dATP-labelled gene probes

1- *Nru*I digest of *H. pylori* UA802
2- *Not*I digest of *H. pylori* UA802
3- *Sfi*I digest of *H. pylori* UA802

Each panel represents an agarose gel on the left and a Southern transfer of that gel, screened with the probe indicated below the panel, on the right. Arrows on the agarose gel show fragments which hybridized. Partial digest fragments are indicated with a "P". PFGE was carried out in 0.75% agarose at 8°C with a constant voltage of 190v, with pulse times of 8h at 10s, 8h at 30s and 8h at 60s for a total running time of 24h. Probes were labelled with α-[^{32}P]-dATP by nick translation. Hybridizations were carried out at 37°C in the presence of 50% formamide.


123 123
123 123
123 123
1 2 3 1 2 3
23S
16S
ureACD
pfr

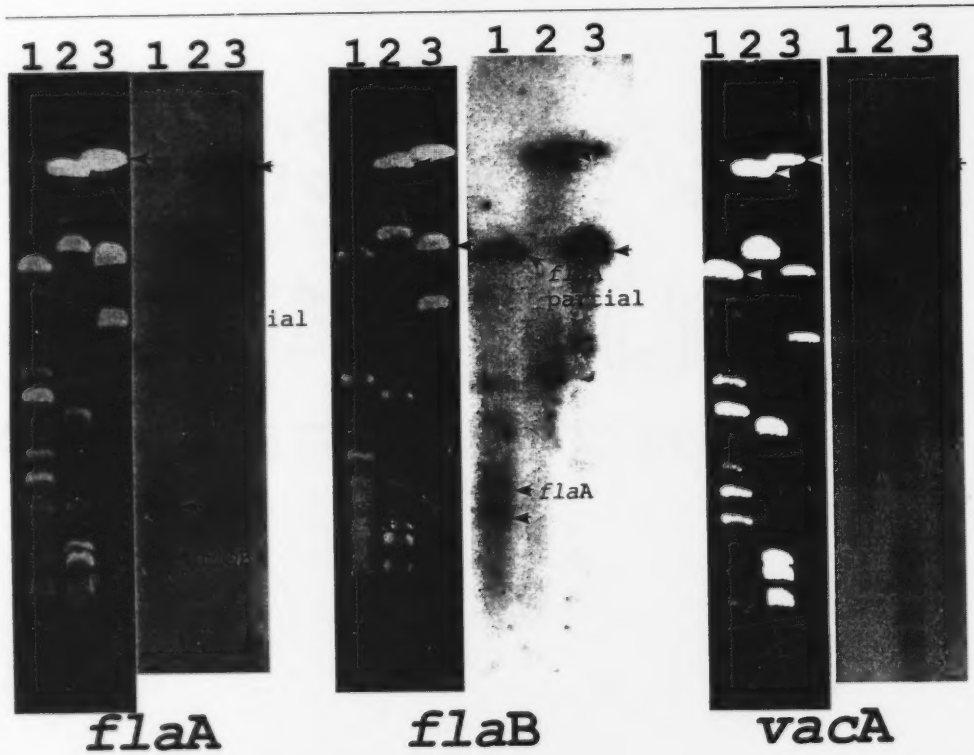

1 2 3 1 2 3
1 2 3
1 2 3
1 2 3 1 2 3
ial
partial
flaA
flaA
flaB
vacA

Table 5-4. Localization of genes to the *H. pylori* UA802 genomic map

Gene	*Nru*I	*Not*I	*Sfi*I
23S rRNA	1,2,4,9	1,3	1,3
16S rRNA	2,3	1,2	1
*Ure*ACD	2	1	2
26 kDa-protein	2	1	1,2
*vac*A	1 (4)	1 (3)	1
pfr	1	2	1
*fla*A	6	2	1
*fla*B	7	1	2
*hpCop*A[a]	3	2	ND
*sod*B	4	7	3
*cag*A	6	2	1
*lpp*20	2	1	ND
*gyr*B	3	2	1
*kat*A[a]	1	6	ND
*hpa*A[a]	1	4	ND

ND- Not done
a- assigned to these fragments by Jiang, 1994

fragment 7 and *Not*I fragment 1. The *gyr*B gene, which did not appear on the original map, was added to the revised map. The data used for the revisions of the *H. pylori* UA802 genomic map were presented to and accepted by the authors of the original map.

Alignment of *H. pylori* genomic maps

In order to compare the arrangements of the genes in different strains of *H. pylori*, the genomic map of *H. pylori* UA802 was aligned with genomic restriction maps constructed by Jiang (1994) of *H. pylori* strains UA861 and NCTC11639 and with the physical map of strain NCTC11638 produced by Bukanov and Berg (1994)(Figure 5-5). The maps were aligned with the *vac*A gene as the starting point.

D. Discussion

D.1. PFGE analysis and construction of a genomic restriction map of *H. pylori* UA802

The genomic restriction map of *H. pylori* UA802 constructed for this research differed from the previously reported map (Taylor *et al.*, 1992a), although the PFGE patterns obtained were identical. Both the arrangement and number of the restriction fragments and the arrangements of genes have been changed. Partial digestions, difficulties in isolating restriction fragments for probes, and inconsistencies in DNA isolation for PFGE have hampered studies of *H. pylori* UA802 and are likely causes for the differences observed. Improvements in preparation of embedded DNA for PFGE, with

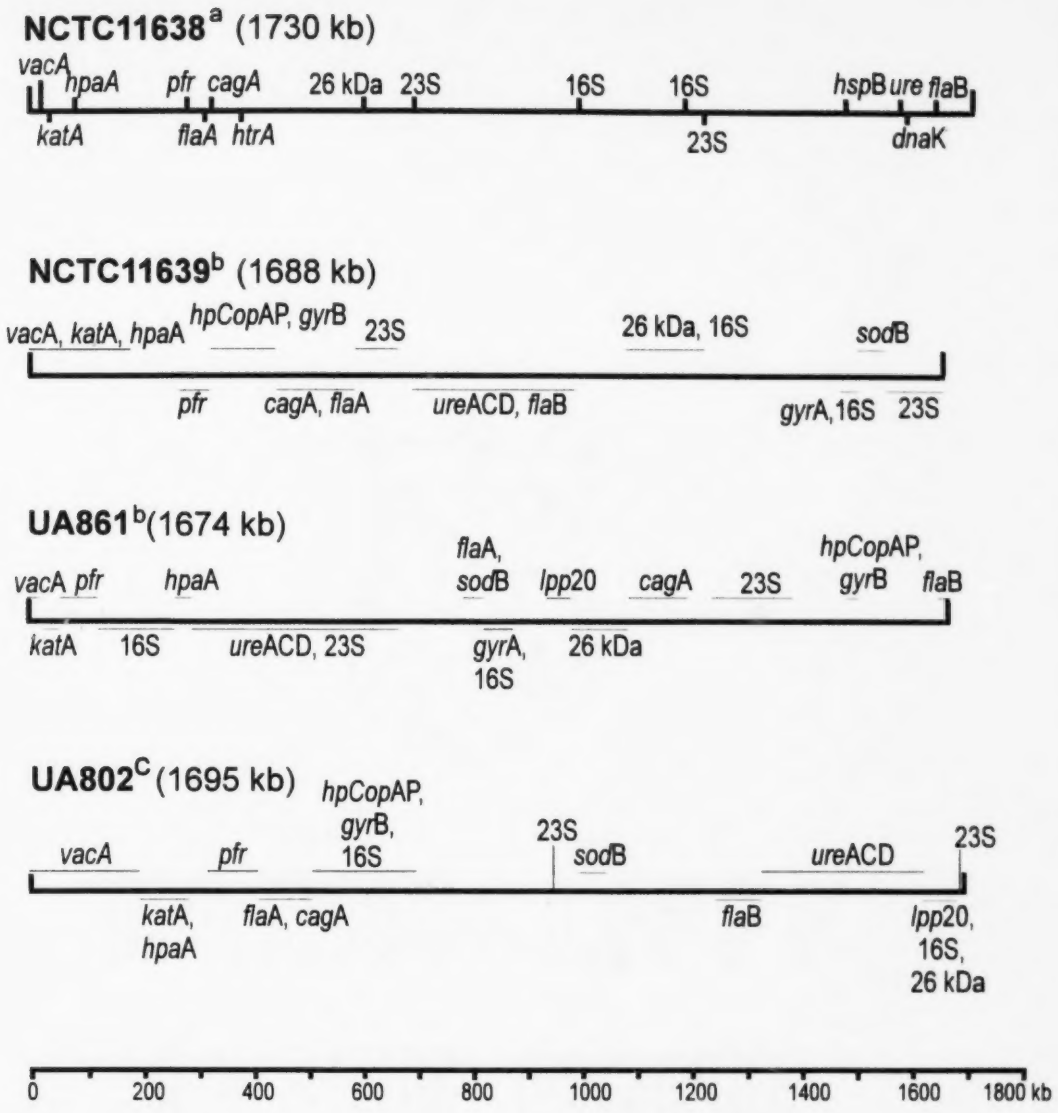

NCTC11638[a] (1730 kb)
vacA hpaA pfr cagA 26 kDa 23S 16S 16S hspB ure flaB
katA flaA htrA 23S dnaK
NCTC11639[b] (1688 kb)
vacA, katA, hpaA hpCopAP, gyrB 23S 26 kDa, 16S sodB
pfr cagA, flaA ureACD, flaB gyrA, 16S 23S
UA861[b](1674 kb)
vacA pfr hpaA flaA, sodB lpp20 cagA 23S hpCopAP, gyrB flaB
katA 16S ureACD, 23S gyrA, 16S 26 kDa
UA802[c] (1695 kb)
vacA pfr hpCopAP, gyrB, 16S 23S sodB ureACD 23S
katA, hpaA flaA, cagA flaB lpp20, 16S, 26 kDa
0 200 400 600 800 1000 1200 1400 1600 1800 kb

Figure 5-5. Alignment of genomic maps of *H. pylori* strains NCTC11638, NCTC11639, UA861 and UA802

The *vac*A gene was used as a starting point for the alignment of the maps.
a- map constructed by Bukanov and Berg, 1994
b- map constructed Jiang, 1994
c- map constructed for this work

the addition of higher EDTA concentrations and the addition of a lysozyme step, as well as refinements in conditions for PFGE (adjustments of pulse times, etc.), have improved the resolution and detection of restriction fragments. This has made the construction of the revised map possible. The additional data obtained with the *Sfi*I restriction digestions have also helped to confirm some of the revisions.

D.2. Comparisons of genome arrangements

Several genes were localized on the genomic restriction map of *H. pylori* UA802. These genes, listed on Table 5-4, represent the majority of genes which have been cloned and characterized from *H. pylori*. Not surprisingly, most of these genes are thought to play a role in the pathogenesis of *H. pylori*. Urease (*ure*ACD) and flagella (*fla*A and *fla*B) have been shown to be important colonization factors (Eaton *et al.*, 1989 and 1991). Superoxide dismutase (*sod*B; Spiegelhalder *et al.*, 1993) and catalase (*kat*A; Lior & Johnson, 1985) could be involved in resistance to phagocytic killing (Cover and Blaser, 1995). The cytotoxin encoded by the *vac*A gene causes damage to epithelial cells and, along with the cytotoxin-associated gene *cag*A, has been associated with the development of ulcers. The rRNA genes (16S and 23S) and gyrase genes (*gyr*A and *gyr*B), while not directly involved in pathogenesis, may be important in specifying antibiotic sensitivities or resistances.

On aligning the genomic map of *H. pylori* UA802 with maps of other *H. pylori* strains, considerable variability in gene

arrangement was observed (Figure 5-5). However, some genes were observed to be loosely associated in all the strains. The gyrase gene, *gyr*B, and the P-type ATPase genes, *hpCop*A and *hpCop*P, localized to the same regions on all of the strains for which they were mapped. The *vac*A, *kat*A, *hpa*A, and *pfr* also appeared to be clustered together in all of the strains. With the exception of strain UA861, *cag*A and *fla*A genes also appeared to have similar arrangements in the genomes. It would appear from these clustered genes that certain regions may be conserved or at least are less susceptible to rearrangement. Krawec and Riley (1990) suggest that nonrandom arrangements, or clustering, of genes suggests that the function or regulation of the genes may be affected by their organization. Since many of the virulence factors of *H. pylori*, such as *vac*A and *cag*A, are variably expressed from strain to strain, it is possible that their arrangement in the chromosome may be a determining factor as to whether they are expressed or not.

Variability has also been observed in restriction digestion patterns of genomic DNA of various strains of *H. pylori* using conventional agarose gels (Majewski *et al.*, 1988; Clayton *et al.*, 1989) and PFGE analyses (Taylor *et al.*, 1992), as well as by arbitrary PCR fingerprinting (Akopyantz *et al.*, 1992). A number of hypotheses for this variability were mentioned in Chapter 1, including differences in methylation patterns, a high frequency of silent mutations, genetic uptake or exchange of DNA and recombination with other strains of *H.*

pylori or other bacterial species, global and/or local rearrangements of the chromosome, as well as environmental pressures (host factors). The wide variation of gene arrangements would support the hypothesis that the genomic diversity was the result of chromosomal rearrangements. The mechanisms by which such rearrangements occur are unclear. Through the course of this study, the *H. pylori* strains used were passaged several times without any observed alteration in PFGE patterns, indicating that such rearrangements were not triggered simply by passaging the organism. Of course, the *in vitro* growth environment differs considerably from *in vivo* conditions where host factors and other factors could trigger rearrangements.

It is unlikely that genomic rearrangement is the sole reason for genomic variability. Although variation in restriction endonuclease profiles would be expected with genomic rearrangement, the extent of variability observed also supports the hypothesis that part of the genomic variability is due to spontaneous mutation and/or genetic exchange (Majewski and Goodwin, 1988). Considerable variations in PCR-based fingerprint analyses further support this hypothesis (Foxall *et al.*, 1992; Fujimoto *et al.*, 1994; Akopyanz *et al.*, 1992a and 1992b). Because *H. pylori* inhabits an environment which is relatively free of other organisms, the possibility of genetic exchange with other organisms is low. However, Akopyants *et al.* (1995) have demonstrated the ability of two strains of *H. pylori* to co-

colonize the stomachs of gnotobiotic piglets, thus creating the possibility for intra-species exchange of DNA. Most likely, a combination of genomic rearrangements, spontaneous mutations and genetic exchange contribute to the genomic variability observed among *H. pylori* strains. Further research is needed to elucidate the exact mechanisms by which genomic variability occurs in *H. pylori*.

Chapter 6

Campylobacter jejuni outer membrane proteins

A. Introduction

Campylobacter species-specific DNA probes were developed in 1990 in this lab (Taylor and Hiratsuka, 1990). One probe, pDT1719, specifically hybridized to DNA from *C. jejuni* and *C. coli*. When a lower hybridization temperature was used to reduce stringency (37°C instead of 42°C), this probe was also able to detect *Campylobacter lari* DNA. The second probe, pDT1720, could only detect *C. jejuni* DNA. Both probes were initially isolated from a λ-gt11 expression library of *C. jejuni* UA580. This library was screened with antiserum directed against the *C. jejuni* UA580 major outer membrane protein (MOMP). Two positive lambda clones were isolated and the *C. jejuni* UA580 DNA fragments from these two clones were further subcloned into pUC13. These plasmid sub-clones, pDT1719, and pDT1720, were used as probes.

To date no genetic characterization of the MOMP of *C. jejuni* has been published although some characterization of the MOMP protein of *C. jejuni* has been reported (Huyer *et al.*,1986; Page *et al.*, 1989). The porin protein has a molecular weight of approximately 45 kDa and was shown to form cation-selective pores in liposome vesicles that allow passage of molecules smaller than ~360 Da.

Since the two cloned DNA probes were isolated from a λ-gt11 expression library using anti-sera directed against

the MOMP, it was possible that all or part of the MOMP gene of this organism had been cloned in these two DNA probes. However, results of characterization of the DNA fragments cloned in these probes, described in this chapter, revealed that the MOMP gene had not been cloned. In this chapter the analysis of the nucleotide sequence of the cloned DNA is described. The N-terminal amino acid sequence for purified MOMP protein was also obtained.

B. Materials and Methods

Please refer to Chapter 2 for detailed methodology used for this chapter.

C. Results

C.1. DNA sequence analysis of pDT1719 and pDT1720

Nucleotide sequencing was carried out to determine if the DNA cloned in the two *Campylobacter* DNA probes encoded all or part of the MOMP of *C. jejuni* UA580. Restriction maps produced during the initial characterization of the two *Campylobacter* probes, pDT1719 and pDT1720, are shown in Figure 6-1. Utilizing the restriction sites available in these plasmids, restriction fragments were sub-cloned into pUC13 and sequenced directly by double stranded DNA sequencing (Wang, 1988) using universal forward and reverse primers. Figure 6-1 shows the direction and length of sequence obtained using these sub-clones. The complete sequence of the *C. jejuni* UA580 DNA fragments cloned in

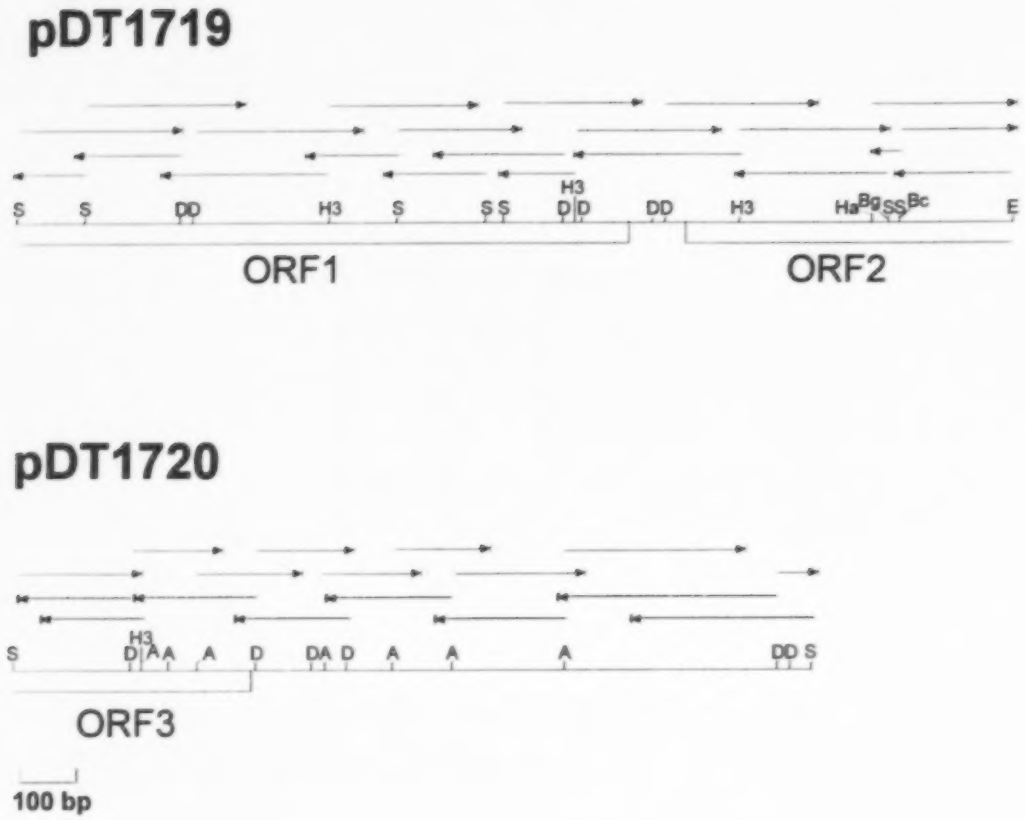



Figure 6-1. Restriction maps of pDT1719 and pDT1720, and sequencing strategy used to determine the sequence of the cloned DNA

A- *Alu*I.
Bc- *Bcl*I
Bg- *Bgl*II
D- *Dra*I
H3- *Hind*III
Ha- *Hae*II
S- *Sau*3AI
Arrows represent size and direction of segments sequenced. Open boxes indicate position of the open reading frames (ORF).

pDT1719 and pDT1720 were obtained. Figure 6-2 shows the DNA sequence for the pDT1719-cloned DNA and Figure 6-3 shows the sequence for pDT1720-cloned DNA. Neither plasmid was found to encode an entire open reading frame (ORF). Instead, pDT1719 was found to encode two partial ORFs, while pDT1720 encoded one partial ORF. The two ORFs found in pDT1719, designated ORF1 and ORF2, were separated by 78 bp. A ribosome binding site was found in this intervening sequence immediately 5' to ORF2 (Figure 6-2). The ORF found in pDT1720, ORF3, encoded only the C-terminal portion of a polypeptide (Figure 6-3).

The DNA and predicted amino acid sequences of the three ORFs were compared to sequences found in the computer database GenBank to find homologous structures in other organisms. No sequences were found which had homology with ORF3 for either the DNA or predicted amino acid sequence. A partial DNA sequence (150 bp) of a putative *C. jejuni* serine protease appeared in GenBank as a direct submission (accession number X82628) which has 97% homology with ORF1. The deduced amino acid sequence of ORF1 also has homology to the HtrA serine protease/heat shock proteins of *Salmonella typhimurium* and *E. coli* (Figure 6-4). ORF1 has 63% and 64% homology respectively with the HtrA proteins of *S. typhimurium* and *E. coli*. Because of this homology, ORF1 was designated *C. jejuni htr*A.

ORF2 shows homology with regulatory-signalling elements of a number of procaryotic two-component regulatory systems

Figure 6-2. The complete nucleotide sequence of the *C.jejuni* DNA cloned in the *Campylobacter* DNA probe pDT1719 and the predicted amino acid sequences encoded by the DNA. The DNA sequence has been submitted to GenBank: accession number U27271

Potential ribosome binding site at position 1176 is underlined and highlighted. Asterisk indicates termination codon.

Figure 6-2

```
       10                  30                  50
GATCAGGGGTGATTATTTCAAAAGATGGTTATATAGTAACAAATAATCACGTTGTAGATG
  S  G  V  I  I  S  K  D  G  Y  I  V  T  N  N  H  V  V  D  D

       70                  90                  110
ATGCTGATACGATTACAGTGAATTTACCAGGAAGCGACATAGAATATAAAGCAAAACTTA
  A  D  T  I  T  V  N  L  P  G  S  D  I  E  Y  K  A  K  L  I

      130                 150                 170
TAGGTAAAGATCCAAAAACAGATTTGGCTGTTATAAAAATAGAGGCTAATAATCTTTCAG
  G  K  D  P  K  T  D  L  A  V  I  K  I  E  A  N  N  L  S  A

      190                 210                 230
CTATTACTTTTACAAATTCTGATGATTTAATGGAAGGAGATGTTGTTTTTGCACTTGGAA
  I  T  F  T  N  S  D  D  L  M  E  G  D  V  V  F  A  L  G  N

      250                 270                 290
ATCCTTTTGGAGTTGGTTTTAGTGTTACAAGTGGGATAATATCTGCTTTAAATAAAGACA
  P  F  G  V  G  F  S  V  T  S  G  I  I  S  A  L  N  K  D  N

      310                 330                 350
ATATAGGTTTAAATCAATATGAAAATTTTATACAAACAGATGCTTCTATCAATCCAGGAA
  I  G  L  N  Q  Y  E  N  F  I  Q  T  D  A  S  I  N  P  G  N

      370                 390                 410
ATTCAGGTGGAGCTTTGGTGGATAGTCGCGGATATTTAGTAGGTATTAATTCAGCTATTC
  S  G  G  A  L  V  D  S  R  G  Y  L  V  G  I  N  S  A  I  L

      430                 450                 470
TTTCTCGTGGTGGTGGAAATAACGGCATAGGTTTTGCCATACCTTCAAATATGGTTAAAG
  S  R  G  G  G  N  N  G  I  G  F  A  I  P  S  N  M  V  K  D

      490                 510                 530
ATATAGCTAAAAAACTTATTGAAAAAGGCAAGATTGATAGAGGATTTTTAGGTGTGACTA
  I  A  K  K  L  I  E  K  G  K  I  D  R  G  F  L  G  V  T  I

      550                 570                 590
TTTTAGCTTTGCAAGGTAATACTAAAAAAGCTTACAAAAATCAAGAAGGAGCTTTAATCA
  L  A  L  Q  G  N  T  K  K  A  Y  K  N  Q  E  G  A  L  I  T

      610                 630                 650
CTGATGTTCAAAAAGGTTCAAGTGCTGATGAAGCAGGGCTTAAGCGTGGAGATTTAGTTA
  D  V  Q  K  G  S  S  A  D  E  A  G  L  K  R  G  D  L  V  T

      670                 690                 710
CTAAAGTTAATGATAAGGTTATAAAAAGTCCTATTGATCTTAAAAATTATATAGGAACTT
  K  V  N  D  K  V  I  K  S  P  I  D  L  K  N  Y  I  G  T  L

      730                 750                 770
TAGAGATTGGTCAAAAAATTTCATTAAGTTACGAAAGAGATGGAGAAAATAAGCAAGCAA
  E  I  G  Q  K  I  S  L  S  Y  E  R  D  G  E  N  K  Q  A  S

      790                 810                 830
GTTTTATTCTTAAAGGCGAAAAAGAAAATCCTAAAGGCGTACAAAGCGATTTGATTGATG
  F  I  L  K  G  E  K  E  N  P  K  G  V  Q  S  D  L  I  D  G

      850                 870                 890
GTTTGAGTTTGAGAAATTTAGATCCAAGACTTAAGGATCGTTTGCAAATTCCAAAAGATG
  L  S  L  R  N  L  D  P  R  L  K  D  R  L  Q  I  P  K  D  V
```

Figure 6-2 continued

```
      910                   930                   950
TTAATGGTGTTTTGGTTCATAGTGTTAAAGAAAAGAGCAAAGGAAAAAATTCTGGTTTCC
  N  G  V  L  V  H  S  V  K  E  K  S  K  G  K  N  S  G  F  Q

      970                   990                  1010
AAGAGGGTGATATTATCATAGGTGTTGGACAAAGTGAAATTAAAAATTTAAAAGATTTAG
  E  G  D  I  I  I  G  V  G  Q  S  E  I  K  N  L  K  D  L  E

     1030                  1050                  1070
AACAAGCTTTAAAACAAGTTAATAAAAAAGAATTTACCAAAGTTTGTGTATATCGCAATG
  Q  A  L  K  Q  V  N  K  K  E  F  T  K  V  C  V  Y  R  N  G

     1090                  1110                  1130
GTTTTGCGACTTTGCTTGTGCTTAAATAAAAATAACGGAAGGATTTTCCTTTCGTTATAA
  F  A  T  L  L  V  L  *

     1150                  1170     RBS          1190
ATTTTAGTTTAAAATTAAAATTAAGTTTAAAATAAAGGAAAACTATGACAAATATTCTTA
                                                M  T  N  I  L  M

     1210                  1230                  1250
TGATAGAAGATGATTTAGAATTAGCAGAAATTACAGCTGAATATTTGGAAAAATTTGATA
  I  E  D  D  L  E  L  A  E  I  T  A  E  Y  L  E  K  F  D  M

     1270                  1290                  1310
TGAAAGTTGATATAGCTCATGAACCTTATATAGGTCTTTCTAAGCTTGCATTAAAAGAAT
  K  V  D  I  A  H  E  P  Y  I  G  L  S  K  L  A  L  K  E  Y

     1330                  1350                  1370
ATCAGCTTATCATTTTAGACCTTTCTTTGCCAGGGCTTGATGGGCTTGAAGTGTGTGAAG
  Q  L  I  I  L  D  L  S  L  P  G  L  D  G  L  E  V  C  E  E

     1390                  1410                  1430
AGATTCGTAAAAAGTATGATACGCCTATTATTGTTTCAAGCGCAAGACATGATATTACAG
  I  R  K  K  Y  D  T  P  I  I  V  S  S  A  R  H  D  I  T  D

     1450                  1470                  1490
ATAAGGTTAATGCTTTAGAGCTTGGAGCAGATGATTACCTACCAAAACCTTATAATCCAA
  K  V  N  A  L  E  L  G  A  D  D  Y  L  P  K  P  Y  N  P  K

     1510                  1530                  1550
AAGAATTACAAGCACGTATTAAAAGTCATTTAAGACGTATTTCAAATACAAAAAGCGCCA
  E  L  Q  A  R  I  K  S  H  L  R  R  I  S  N  T  K  S  A  I

     1570                  1590                  1610
TAGCAAAAAGTGTAAAAGATCTTGTTTATGATCAATATAAGCATATTATTACCATGAAAG
  A  K  S  V  K  D  L  V  Y  D  Q  Y  K  H  I  I  T  M  K  G

     1630                  1650                  1670
GACAAGAGCTTACTTTAACTAATGCTGAATTTGATATTTTAAGCTATTTGATTAAAAAAG
  Q  E  L  T  L  T  N  A  E  F  D  I  L  S  Y  L  I  K  K  E

     1690                  1710                  1730
AGGGTGGAGTGGTAAGTCGTGAAGAGCTTGTTTATAATTGTTCTTCTATTAGTGAGGATT
  G  G  V  V  S  R  E  E  L  V  Y  N  C  S  S  I  S  E  D  S

     1750                  1770
CCAGTAATAAAAGTATAGATGTTATTATCAGAGAATTC
  S  N  K  S  I  D  V  I  I  R  E  F
```

Figure 6-3. The complete nucleotide sequence of the *C.jejuni* DNA cloned in the *Campylobacter* probe pDT1720 and the predicted amino acid sequences encoded by the DNA. The DNA sequence has been submitted to GenBank: accession number U27272. Asterisks indicate termination codons.

Figure 6-3

```
        10                  30                  50
GATCCTAGCAATAAAAATCTTTTAGAACAACTTAAAAATAAAAATACCAATCTTTACGCT
D  P  S  N  K  N  L  L  E  Q  L  K  N  K  N  T  N  L  Y  A

        70                  90                  110
ATCTTTTTACTAAAAGAAAATATTAATGATTTTAATAATACAACTCTCCAAAATGAACTA
I  F  L  L  K  E  N  I  N  D  F  N  N  T  T  L  Q  N  E  L

        130                 150                 170
AAACAAATTTATAATAATGCTCAAACAAACACTTTACTTAAAAATATTATTGCATTATCT
K  Q  I  Y  N  N  A  Q  T  N  T  L  L  K  N  I  I  A  L  S

        190                 210                 230
TTAGGGGATAAATCGATATTTTTAAAAAATTATGATAAATTATTAGAAGCTTACAAACTT
L  G  D  K  S  I  F  L  K  N  Y  D  K  L  L  E  A  Y  K  L

        250                 270                 290
TTAGAGCAAAATAAAATCGAAGAAGCTAATGTTTTACTTTCTCAGATAAAAGAAAATTCA
L  E  Q  N  K  I  E  E  A  N  V  L  L  S  Q  I  K  E  N  S

        310                 330                 350
AGTTTAAATCAAATAGCTAAAAACCTCAACTTTCAGGCATCCATGAAAATTTTCCTATTG
S  L  N  Q  I  A  K  N  L  N  F  Q  A  S  M  K  I  F  L  L

        370                 390                 410
TATTCTTGGCGTGGCCCAAAGCAATATTTGACCAAAACAAACCGATGGAGATTTAAGTCA
Y  S  W  R  G  P  K  Q  Y  L  T  K  T  N  R  W  R  F  K  S

        430                 450                 470
TAATAGCAATTTAAAAAGCAAAATCGTTGATTGGAATCTAGTTTCTGCAAAACTTAGCAA
*  *

        490                 510                 530
TAATACAACTATCTTAAAAAAATAATATCTCCATAGACAAGTTTAAACTTCCTAAAAATT

        550                 570                 590
ATATGCTTCTAGCTTATCAAGATGGGGGATATTTTACTGCAGATAATAATGGAAATTTAA

        610                 630                 650
AAATTTTTGACAATTCTTATAATGAAATTTATAATTTTCAATTTGACGCTAGTGTTGTAG

        670                 690                 710
GAGTAGCTTCTAATGGAGATGATTTAGCCCTTGTACTTGCAAACAATACTATAGTTCTTG

        730                 750                 770
CAAATCGTTCTTTGGGTATTAAATTCAGTCAAACCCTAACCTCAGCTCCAGCCCAAGACA

        790                 810                 830
GTCGTACCGCTAATCCTATATTTTTAGACAATATCATCGTTTATCCAACACTAGATGGTA

        850                 870                 890
AAATTTTAATTCTTTCAAGAAACAATCTTCAAATCATTAAAGATGTGGTTATTTCAGCAG

        910                 930                 950
AAAATTTTTTCAACAATGTAATCCATCTAAGTGTAATCGGAGATAAACTTATAGCAGCAA
```

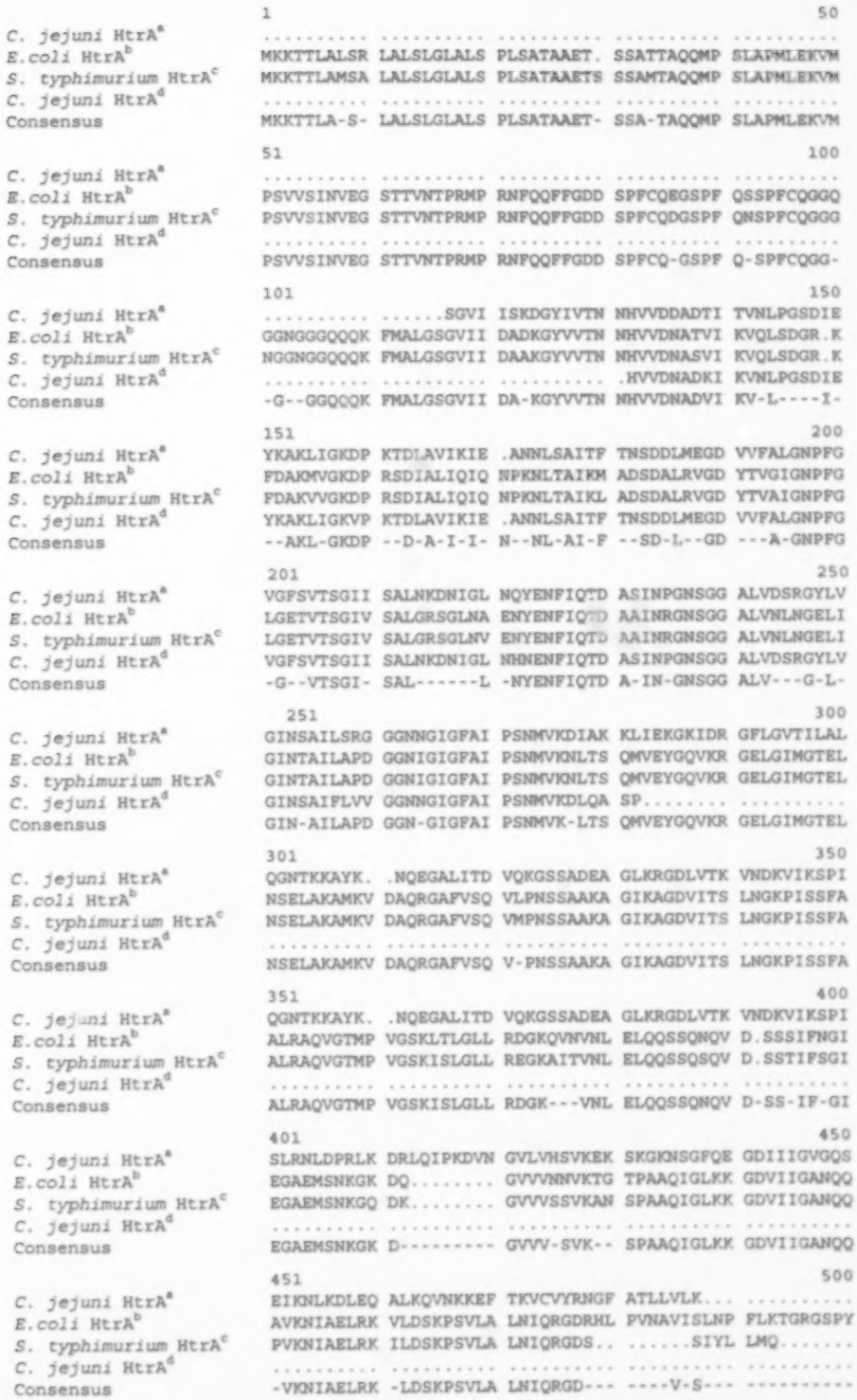
Figure 6-3 continued

```
       970                  990                 1010
CAGCTAAAAAATCATTGTTGTAAGCCCTGCTAGAAACTCTTTATCTTGACGCTGACATAA

      1030                 1050                 1070
AAGATGTTGCACTTAGCGATGATGGTATTTTTATACTAGAAAAAGATGGCACTATTATTA

      1090                 1110                 1130
AAACTGATTATAACTTAAGAAAAATTGCGGAAAAAAAAATTTGAATTTGCTATTTTTGTAA

      1150                 1170                 1190
AAAGCAATATATACAATAATTATTTATATATTTTTGAAAAAACAGGCTATCTTATCAAAA

      1210                 1230                 1250
TGAATCTGAATTTAGATAATGCACAAGTATTTAAACTTTCAGAAGCTGTTGATAAAATTT

      1270                 1290                 1310
CTTTCATGGGAAATGGTAAATTTTACTACGGAGATAAAATTTTAGATTTGCTATGATAAA

      1330                 1350                 1370
AATTTTAGAACAAACAATAAAGGCTTTAAAATTAAATTTAAAGCCCTATGACTTAAGTAT

      1390                 1410
GTTGACTAGAAAAAAAAGTTATATTTGTGCTAAAGATC
```

Figure 6-4. Amino acid sequence comparisons of the putative *C. jejuni* HtrA protein with sequences found in GenBank

a- predicted amino acid sequence of the putative *htr*A gene of *C. jejuni* UA580 determined in this research. Submitted to GenBank: Accession # U27271.
b- translation of DNA sequence determined by Lipinska *et al.*, 1988. Accession #X12457
c- translation of DNA sequence from GenBank direct submission 1994. Accession #A18802
d- translation of DNA sequence from GenBank direct submission 1994. Accession #X82628

Figure 6-4


1 50
C. jejuni HtrA[a] ....
E.coli HtrA[b] MKKTTLALSR LALSLGLALS PLSATAAET. SSATTAQQMP SLAPMLEKVM
S. typhimurium HtrA[c] MKKTTLAMSA LALSLGLALS PLSATAAETS SSAMTAQQMP SLAPMLEKVM
C. jejuni HtrA[d] ....
Consensus MKKTTLA-S- LALSLGLALS PLSATAAET- SSA-TAQQMP SLAPMLEKVM
51 100
C. jejuni HtrA[a] ....
E.coli HtrA[b] PSVVSINVEG STTVNTPRMP RNFQQFFGDD SPFCQEGSPF QSSPFCQGGQ
S. typhimurium HtrA[c] PSVVSINVEG STTVNTPRMP RNFQQFFGDD SPFCQDGSPF QNSPFCQGGG
C. jejuni HtrA[d] ....
Consensus PSVVSINVEG STTVNTPRMP RNFQQFFGDD SPFCQ-GSPF Q-SPFCQGG-
101 150
C. jejuni HtrA[a] ....SGVI ISKDGYIVTN NHVVDDADTI TVNLPGSDIE
E.coli HtrA[b] GGNGGGQQQK FMALGSGVII DADKGYVVTN NHVVDNATVI KVQLSDGR.K
S. typhimurium HtrA[c] NGGNGGQQQK FMALGSGVII DAAKGYVVTN NHVVDNASVI KVQLSDGR.K
C. jejuni HtrA[d] .... .HVVDNADKI KVNLPGSDIE
Consensus -G--GGQQQK FMALGSGVII DA-KGYVVTN NHVVDNADVI KV-L----I-
151 200
C. jejuni HtrA[a] YKAKLIGKDP KTDLAVIKIE .ANNLSAITF TNSDDLMEGD VVFALGNPFG
E.coli HtrA[b] FDAKMVGKDP RSDIALIQIQ NPKNLTAIKM ADSDALRVGD YTVGIGNPFG
S. typhimurium HtrA[c] FDAKVVGKDP RSDIALIQIQ NPKNLTAIKL ADSDALRVGD YTVAIGNPFG
C. jejuni HtrA[d] YKAKLIGKVP KTDLAVIKIE .ANNLSAITF TNSDDLMEGD VVFALGNPFG
Consensus --AKL-GKDP --D-A-I-I- N--NL-AI-F --SD-L--GD ---A-GNPFG
201 250
C. jejuni HtrA[a] VGFSVTSGII SALNKDNIGL NQYENFIQTD ASINPGNSGG ALVDSRGYLV
E.coli HtrA[b] LGETVTSGIV SALGRSGLNA ENYENFIQTD AAINRGNSGG ALVNLNGELI
S. typhimurium HtrA[c] LGETVTSGIV SALGRSGLNV ENYENFIQTD AAINRGNSGG ALVNLNGELI
C. jejuni HtrA[d] VGFSVTSGII SALNKDNIGL NHNENFIQTD ASINPGNSGG ALVDSRGYLV
Consensus -G--VTSGI- SAL------L -NYENFIQTD A-IN-GNSGG ALV---G-L-
251 300
C. jejuni HtrA[a] GINSAILSRG GGNNGIGFAI PSNMVKDIAK KLIEKGKIDR GFLGVTILAL
E.coli HtrA[b] GINTAILAPD GGNIGIGFAI PSNMVKNLTS QMVEYGQVKR GELGIMGTEL
S. typhimurium HtrA[c] GINTAILAPD GGNIGIGFAI PSNMVKNLTS QMVEYGQVKR GELGIMGTEL
C. jejuni HtrA[d] GINSAIFLVV GGNNGIGFAI PSNMVKDLQA SP..... .....
Consensus GIN-AILAPD GGN-GIGFAI PSNMVK-LTS QMVEYGQVKR GELGIMGTEL
301 350
C. jejuni HtrA[a] QGNTKKAYK. .NQEGALITD VQKGSSADEA GLKRGDLVTK VNDKVIKSPI
E.coli HtrA[b] NSELAKAMKV DAQRGAFVSQ VLPNSSAAKA GIKAGDVITS LNGKPISSFA
S. typhimurium HtrA[c] NSELAKAMKV DAQRGAFVSQ VMPNSSAAKA GIKAGDVITS LNGKPISSFA
C. jejuni HtrA[d] ....
Consensus NSELAKAMKV DAQRGAFVSQ V-PNSSAAKA GIKAGDVITS LNGKPISSFA
351 400
C. jejuni HtrA[a] QGNTKKAYK. .NQEGALITD VQKGSSADEA GLKRGDLVTK VNDKVIKSPI
E.coli HtrA[b] ALRAQVGTMP VGSKLTLGLL RDGKQVNVNL ELQQSSQNQV D.SSSIFNGI
S. typhimurium HtrA[c] ALRAQVGTMP VGSKISLGLL REGKAITVNL ELQQSSQSQV D.SSTIFSGI
C. jejuni HtrA[d] ....
Consensus ALRAQVGTMP VGSKISLGLL RDGK---VNL ELQQSSQNQV D-SS-IF-GI
401 450
C. jejuni HtrA[a] SLRNLDPRLK DRLQIPKDVN GVLVHSVKEK SKGKNSGFQE GDIIIGVGQS
E.coli HtrA[b] EGAEMSNKGK DQ..... GVVVNNVKTG TPAAQIGLKK GDVIIGANQQ
S. typhimurium HtrA[c] EGAEMSNKGQ DK..... GVVVSSVKAN SPAAQIGLKK GDVIIGANQQ
C. jejuni HtrA[d] ....
Consensus EGAEMSNKGK D--------- GVVV-SVK-- SPAAQIGLKK GDVIIGANQQ
451 500
C. jejuni HtrA[a] EIKNLKDLEQ ALKQVNKKEF TKVCVYRNGF ATLLVLK... .....
E.coli HtrA[b] AVKNIAELRK VLDSKPSVLA LNIQRGDRHL PVNAVISLNP FLKTGRGSPY
S. typhimurium HtrA[c] PVKNIAELRK ILDSKPSVLA LNIQRGDS.. ......SIYL LMQ.....
C. jejuni HtrA[d] ....
Consensus -VKNIAELRK -LDSKPSVLA LNIQRGD--- ----V-S--- ----------

(Figure 6-5). Because ORF2 was found adjacent to the *C. jejuni htr*A, and appears to be expressed with *htr*A in an operon, it was designated *htr*B.

C.2. N-terminal amino acid sequencing of the MOMP of *C. jejuni* UA580

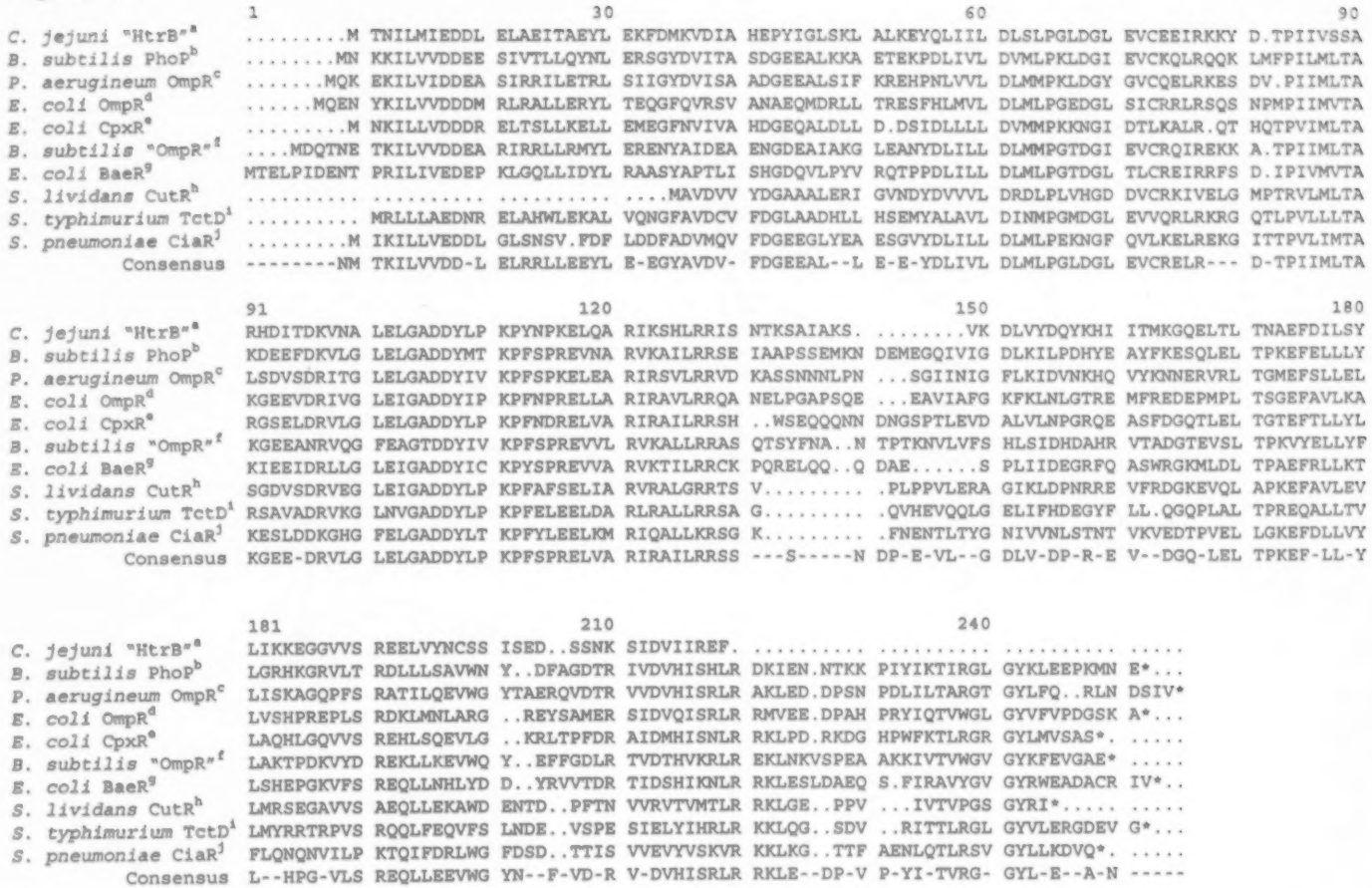
By obtaining the N-terminal sequence of the purified MOMP protein, it would be possible to design oligonucleotide probes to identify clones which carry the MOMP gene. Thus, SDS-PAGE of MOMP preparations and Western blots of these gels to PVDF membranes were carried out for N-terminal sequence analysis. SDS-PAGE revealed that the MOMP preparations were not pure (Figure 6-6). These were the same MOMP preparations used to inoculate rabbits in the preparation of antiserum against the MOMP protein (Taylor and Hiratsuka, 1990) which was subsequently used for the screening of a λ-gt11 library of *C. jejuni* UA580 DNA. The MOMP band at approximately 45 kDa was excised from Western blots of the protein on PVDF membranes for N-terminal sequencing analysis. The first 30 amino acids of the MOMP protein were determined (Figure 6-7). This amino acid sequence was compared with those found in the SwissPro protein database and was found to have 59% identity with the N-terminal sequence of two porin proteins from the organism *Campylobacter rectus*, formerly *Wolinella recta* (Figure 6-7).

The amino-acid sequence was intended to be used for designing oligonucleotide probes for detection of MOMP clones of *C. jejuni*. However, the focus of my research was shifted

Figure 6-5. Amino acid sequence comparisons of the putative *C. jejuni* HtrB protein with sequences found in GenBank

a- *Campylobacter jejuni* "HtrB". Translation of the DNA sequence of ORF2 of plasmid pDT1719. This thesis. Accession #U27271
b- *Bacillus subtilis* PhoP. Translation of DNA sequence determined by Lee and Hullet, 1992. Accession #X67676
c- *Porphyridium aerugineum* OmpR. Translation of DNA sequence determined by Kessler *et al.*, 1992. Accession #X62579
d- *Escherichia coli* OmpR. Translation of DNA sequence determined by Wurtzel *et al.*, 1982. Accession #J01656
e- *E. coli* CpxR. Translation of DNA sequence determined by Dong *et al.*, 1993. Accession #L14579
f- *B. subtilis* "OmpR". Translation of DNA sequence determined by Yamamoto *et al.*, 1991. Accession #L09228
g- *E. coli* BaeR. Translation of DNA sequence determined by Nagasawa *et al.*, 1993. Accession #D14054
h- *Streptomyces lividans* CutR. Translation of DNA sequence determined by Tseng and Chen, 1991. Accession #S67525
i- *Salmonella typhimurium* TctD. Translation of DNA sequence determined by Widenhorn, *et al.*, 1989. Accession #28368
j- *Streptococcus pneumoniae* CiaR. Translation of DNA sequence determined by Guenzi, *et al.*, 1994. Accession #X77249

Figure 6-5


1 30 60 90
C. jejuni "HtrB"[a] .....M TNILMIEDDL ELAEITAEYL EKFDMKVDIA HEPYIGLSKL ALKEYQLIIL DLSLPGLDGL EVCEEIRKKY D.TPIIVSSA
B. subtilis PhoP[b] .....MN KKILVVDDEE SIVTLLQYNL ERSGYDVITA SDGEEALKKA ETEKPDLIVL DVMLPKLDGI EVCKQLRQQK LMFPILMLTA
P. aerugineum OmpR[c] .....MQK EKILVIDDEA SIRRILETRL SIIGYDVISA ADGEEALSIF KREHPNLVVL DLMMPKLDGY GVCQELRKES DV.PIIMLTA
E. coli OmpR[d] .....MQEN YKILVVDDDM RLRALLERYL TEQGFQVRSV ANAEQMDRLL TRESFHLMVL DLMLPGEDGL SICRRLRSQS NPMPIIMVTA
E. coli CpxR[e] .....M NKILLVDDDR ELTSLLKELL EMEGFNVIVA HDGEQALDLL D.DSIDLLLL DVMMPKKNGI DTLKALR.QT HQTPVIMLTA
B. subtilis "OmpR"[f] ....MDQTNE TKILVVDDEA RIRRLLRMYL ERENYAIDEA ENGDEAIAKG LEANYDLILL DLMMPGTDGI EVCRQIREKK A.TPIIMLTA
E. coli BaeR[g] MTELPIDENT PRILIVEDEP KLGQLLIDYL RAASYAPTLI SHGDQVLPYV RQTPPDLILL DLMLPGTDGL TLCREIRRFS D.IPIVMVTA
S. lividans CutR[h] ....MAVDVV YDGAAALERI GVNDYDVVVL DRDLPLVHGD DVCRKIVELG MPTRVLMLTA
S. typhimurium TctD[i] ..... MRLLLAEDNR ELAHWLEKAL VQNGFAVDCV FDGLAADHLL HSEMYALAVL DINMPGMDGL EVVQRLRKRG QTLPVLLLTA
S. pneumoniae CiaR[j] .....M IKILLVEDDL GLSNSV.FDF LDDFADVMQV FDGEEGLYEA ESGVYDLILL DLMLPEKNGF QVLKELREKG ITTPVLIMTA
Consensus --------NM TKILVVDD-L ELRRLLEEYL E-EGYAVDV- FDGEEAL--L E-E-YDLIVL DLMLPGLDGL EVCRELR--- D-TPIIMLTA
91 120 150 180
C. jejuni "HtrB"[a] RHDITDKVNA LELGADDYLP KPYNPKELQA RIKSHLRRIS NTKSAIAKS. .....VK DLVYDQYKHI ITMKGQELTL TNAEFDILSY
B. subtilis PhoP[b] KDEEFDKVLG LELGADDYMT KPFSPREVNA RVKAILRRSE IAAPSSEMKN DEMEGQIVIG DLKILPDHYE AYFKESQLEL TPKEFELLLY
P. aerugineum OmpR[c] LSDVSDRITG LELGADDYIV KPFSPKELEA RIRSVLRRVD KASSNNNLPN ...SGIINIG FLKIDVNKHQ VYKNNERVRL TGMEFSLLEL
E. coli OmpR[d] KGEEVDRIVG LEIGADDYIP KPFNPRELLA RIRAVLRRQA NELPGAPSQE ...EAVIAFG KFKLNLGTRE MFREDEPMPL TSGEFAVLKA
E. coli CpxR[e] RGSELDRVLG LELGADDYLP KPFNDRELVA RIRAILRRSH ..WSEQQQNN DNGSPTLEVD ALVLNPGRQE ASFDGQTLEL TGTEFTLLYL
B. subtilis "OmpR"[f] KGEEANRVQG FEAGTDDYIV KPFSPREVVL RVKALLRRAS QTSYFNA..N TPTKNVLVFS HLSIDHDAHR VTADGTEVSL TPKVYELLYF
E. coli BaeR[g] KIEEIDRLLG LEIGADDYIC KPYSPREVVA RVKTILRRCK PQRELQQ..Q DAE.....S PLIIDEGRFQ ASWRGKMLDL TPAEFRLLKT
S. lividans CutR[h] SGDVSDRVEG LEIGADDYLP KPFAFSELIA RVRALGRRTS V..... .PLPPVLERA GIKLDPNRRE VFRDGKEVQL APKEFAVLEV
S. typhimurium TctD[i] RSAVADRVKG LNVGADDYLP KPFELEELDA RLRALLRRSA G..... .QVHEVQQLG ELIFHDEGYF LL.QGQPLAL TPREQALLTV
S. pneumoniae CiaR[j] KESLDDKGHG FELGADDYLT KPFYLEELKM RIQALLKRSG K..... .FNENTLTYG NIVVNLSTNT VKVEDTPVEL LGKEFDLLVY
Consensus KGEE-DRVLG LELGADDYLP KPFSPRELVA RIRAILRRSS ---S-----N DP-E-VL--G DLV-DP-R-E V--DGQ-LEL TPKEF-LL-Y
181 210 240
C. jejuni "HtrB"[a] LIKKEGGVVS REELVYNCSS ISED..SSNK SIDVIIREF. ....
B. subtilis PhoP[b] LGRHKGRVLT RDLLLSAVWN Y..DFAGDTR IVDVHISHLR DKIEN.NTKK PIYIKTIRGL GYKLEEPKMN E*...
P. aerugineum OmpR[c] LISKAGQPFS RATILQEVWG YTAERQVDTR VVDVHISRLR AKLED.DPSN PDLILTARGT GYLFQ..RLN DSIV*
E. coli OmpR[d] LVSHPREPLS RDKLMNLARG ..REYSAMER SIDVQISRLR RMVEE.DPAH PRYIQTVWGL GYVFVPDGSK A*...
E. coli CpxR[e] LAQHLGQVVS REHLSQEVLG ..KRLTPFDR AIDMHISNLR RKLPD.RKDG HPWFKTLRGR GYLMVSAS*. .....
B. subtilis "OmpR"[f] LAKTPDKVYD REKLLKEVWQ Y..EFFGDLR TVDTHVKRLR EKLNKVSPEA AKKIVTVWGV GYKFEVGAE* .....
E. coli BaeR[g] LSHEPGKVFS REQLLNHLYD D..YRVVTDR TIDSHIKNLR RKLESLDAEQ S.FIRAVYGV GYRWEADACR IV*..
S. lividans CutR[h] LMRSEGAVVS AEQLLEKAWD ENTD..PFTN VVRVTVMTLR RKLGE..PPV ...IVTVPGS GYRI*..... .....
S. typhimurium TctD[i] LMYRRTRPVS RQQLFEQVFS LNDE..VSPE SIELYIHRLR KKLQG..SDV ..RITTLRGL GYVLERGDEV G*...
S. pneumoniae CiaR[j] FLQNQNVILP KTQIFDRLWG FDSD..TTIS VVEVYVSKVR KKLKG..TTF AENLQTLRSV GYLLKDVQ*. .....
Consensus L--HPG-VLS REQLLEEVWG YN--F-VD-R V-DVHISRLR RKLE--DP-V P-YI-TVRG- GYL-E--A-N -----

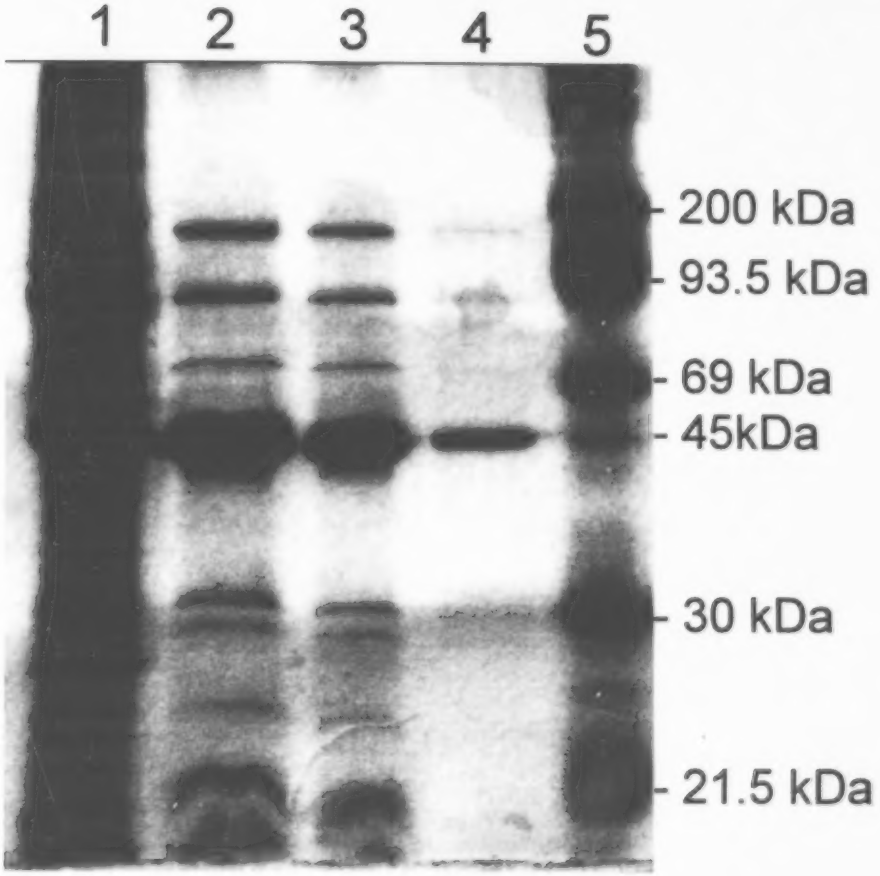


Figure 6-6. SDS-PAGE of purified MOMP samples from *C. jejuni* UA580

Lane 1- *C. jejuni* UA580 sonicate
Lane 2- 10 µl of MOMP preparation
Lane 3- 5 µl of MOMP preparation
Lane 4- 1 µl of MOMP preparation
Lane 5- molecular weight marker

	1	30
C. rectus 43 kDa OMP	TPLEEAIKDV	DFSGFAAYXY TGNKLXVNNK
C. rectus 51 kDa OMP	TPLEEAIKNV	DLSGYAXYXY NNITVKKNPG
C. jejuni 45 kDa MOMP	TPLEEAIKDV	DVSGL.RYRY DTGNFDKNFV
Consensus	**TPLEEAIK-V**	**D-SG-ARY-Y --GT-DKN--**

Figure 6-7. N-terminal amino acid sequence of the *C. jejuni* UA580 45 kDa MOMP and comparison with the N-terminal sequence of the *C. rectus* 43 kDa and 51 kDa membrane proteins

The *C. rectus* sequences were determined by Kennell and Holt, 1991.

to work on *H. pylori* and no further research was conducted on the *C. jejuni* MOMP.

D. Discussion

D.1. Characterization of pDT1719 and pDT1720

DNA sequencing of the *C. jejuni* UA580 DNA fragments cloned in pDT1719 and pDT1720 was carried out for this research. Amino acid and DNA homology analyses were carried out on the DNA sequences and deduced amino acid sequences of the three ORFs encoded by pDT1719 and pDT1720 against the GenBank database. Because of strong homology observed with the *htr*A genes of *E. coli* and *S. typhimurium*, ORF1 was designated *htr*A. In *E. coli* this protein is expressed in response to environmental stress and is necessary for survival at elevated temperatures (Lipinska *et al.*, 1989). The exact role of these proteins is not clear but they are believed to function as proteases which specifically degrade protein by-products and abnormal proteins which accumulate in the periplasm or inner membrane during heat shock (Lipinska *et al.*, 1990). In *E. coli*, expression of the *htr*A gene is controlled by the σ^E protein (Erickson and Gross, 1989). Because the 5' end of the ORF1 sequence was not cloned, the control of expression of the putative HtrA protein in *C. jejuni* remains unclear.

ORF2, which was designated *htr*B since it appeared to be linked to *htr*A in an operon, showed homology with several response-regulatory elements involved in two-component

regulatory systems. In these two-component regulatory systems, one component, the sensory receptor, responds to environmental factors, activating the second component, the response-regulatory element (Stock *et al.*, 1990). In most instances, the sensory receptor is a protein kinase which first activates itself by auto-phosphorylation, then transfers the phosphoryl group to the response-regulatory element thereby activating it. This response-regulatory element then goes on to specifically activate or repress transcription of a gene or set of genes to respond to the environmental conditions (Stock *et al.*, 1990). The environmental factors which lead to the activation of the putative regulatory element I have cloned, *htr*B, could not be determined. Since *htr*B was linked to *htr*A, a serine protease/heat shock protein, it was possible that *htr*B was involved in the heat shock response of *C. jejuni*. In addition, since these two genes appear to be linked together, it would be interesting to determine if and how their products interact. The HtrA protein may have kinase activity and may activate the HtrB protein in a two-component regulatory system responding to heat shock. It is also possible that the HtrA protein activates the HtrB protein by specific protease activity. A third possibility is that the HtrB protein is activated by a protein other than HtrA. Further research is needed to answer these questions.

Since the preparation of MOMP used to raise antibodies was shown to contain other proteins (Figure 6-6), the HtrA or B

proteins, as well as the protein encoded by ORF3, were likely present in the "MOMP" preparation, and antibodies against these proteins were raised. This would account for their detection and isolation from the λ-gt11 expression library. Since porins are identified by their secondary and tertiary amino acid structure, even though ORF3 did not show homology to any DNA or protein sequence in the GenBank database, it was still possible that it represented the C-terminal portion of the MOMP protein. If this was the case, then the complete sequence of this gene would have to be determined and the three-dimensional structure of the protein it encodes would identify it as a porin.

D.2. Comparison of the N-terminal sequence of the *C. jejuni* MOMP to the N-terminal sequence of the porin protein of *Campylobacter rectus*

The results of the DNA sequencing of the two *Campylobacter* probes revealed that no complete open reading frames had been cloned. The sequence also showed that none of ORFs had homology with any know porin proteins. It was possible that the MOMP protein could not be expressed in the expression vector system and perhaps a different strategy was needed to identify MOMP clones. By determining the N-terminal amino acid sequence of the MOMP protein of *C. jejuni* UA580, oligonucleotide probes could be designed based on this sequence and used for identification of MOMP clones.

The first thirty amino acids of the N-terminal amino acid sequence of the 45 kDa MOMP of *C. jejuni* UA580 were determined and represented the first reported sequence for this protein. This amino acid sequence was found to have extensive homology with two membrane proteins (45 and 51 kDa) of *C. rectus* (Kennell and Holt, 1991). The 51 kDa protein was shown to be a porin, while the function of the 45 kDa protein could not be determined. The MOMP of *C. jejuni* has also been demonstrated to be a porin (Huyer *et al.*, 1986). The extensive homology observed between the N-terminal amino acid sequences of the outer membrane proteins of *C. rectus* and *C. jejuni* was consistent with the observed similarities in rRNA sequences between *C. rectus* and other *Campylobacter* species and supports the reclassification of *C. rectus* from its former genus, *Wolinella*, to the genus *Campylobacter*. This comparison demonstrates the possible use of the MOMP N-terminal amino acid sequence as a taxonomic marker for the genus *Campylobacter*. If the gene for this MOMP could be cloned in the future, all or part of the gene may be useful as a genus-specific probe for *Campylobacter* species.

Chapter 7

General Discussion

A. *Helicobacter pylori* genetics

A.1. Genomic libraries of *H. pylori* UA802

An objective of this research was to construct genomic DNA libraries of *H. pylori* UA802 DNA. This was accomplished using the lambda cloning vector λ-DASH. Less efficient libraries were also constructed with other vectors (pUC20, pHC79, pSa747, and λ-GEM11). With all of these different vectors, restriction modification was one of the main barriers which had to be overcome to enable efficient library construction. The λ-DASH libraries were constructed using the host strain SRB which, although highly mutated for restriction modification functions, had been found to be less efficient than other strains, such as *E. coli* ER1792 (Phadnis *et al.*, 1993). *E. coli* ER1792 and *E. coli* SRB share the same restriction modification phenotype, but Phadnis *et al.* showed *E. coli* ER1792 to be almost one hundred times more efficient than SRB as an *H. pylori* lambda host. Both strains carry deletions of the *mcr*C-*hsd*RMS-*mmr* region, but at different points. Phadnis *et al.* speculated that another restriction modification locus specific for *H. pylori* modified DNA may have been affected by the deletion in *E. coli* ER1792 but not in *E. coli* SRB, or that *E. coli* ER1792 had an *H. pylori*-specific modification function not found in *E. coli* SRB. The combined effects of a number of factors, such as sequence-

related factors (*E. coli* promoter-like sequences), and lethal genetic elements, in addition to restriction modification, may also have contributed to the low efficiencies observed in library construction (Clayton *et al.*, 1989a; Morrison and Jaurin, 1990; Covacci *et al.*, 1993).

My results show that the libraries were not representative of the entire genome. When screened with either the 16S or 23S probes, 7.5-18% of the plaques were positive. This greatly exceeded the expected 2% value estimated for a representative library. This suggests that these clones must have had some selective advantage over the other clones or that conditions selected for these clones or selected against other clones. This phenomenon was also observed during the construction of genomic libraries of the nematode *Caenorhabditis elegans* (Ellis *et al.*, 1986; Gibson *et al.*, 1987), where rDNA sequences were cloned at a frequency much higher than expected using both lambda vectors and cosmid vectors. Two possible reasons were advanced for this phenomenon (Gibson *et al.*, 1987). First, since rDNA did not encode a protein, it was less likely to produce toxic products and may have had greater stability. Second, the G + C content of the rDNA (49.3 mol%) was much higher than the rest of the genome (36 mol%), reducing sequence-related instability. Similarly, the DNA sequence of the cloned 23S-5S genes of *H. pylori* UA802 also showed a higher G + C ratio (45.3 mol%) than the ratio calculated for the *H. pylori* genome (35.2 mol%; Béji *et al.*, 1988). Other genes or

regions of the chromosome with higher localized G + C ratios may have a similar advantage over regions with lower G + C and thus may also be represented at a higher than normal frequency. Conversely, regions which have a lower G + C content may be poorly represented in such libraries. The hypothesis that localized G + C ratios can affect the stability of cloned DNA is corroborated by the results of genomic library construction by Bukanov and Berg (1994), who produced a genomic library of *H. pylori* NCTC11368 using the cosmid vector Lorist6. This vector was designed to counteract the effects of sequence-related instabilities associated with AT-rich DNA. With this vector, 16S and 23S genes were cloned at the expected frequencies. For this research screening of the libraries was not carried out on primary plates, but rather on plates made from overlays of the primary plates. It was possible that the frequency of 23S and 16S clones on the primary plates may have been closer to the expected frequency and any advantage that these clones may have had were amplified in the overlay plates.

Other researchers have shown that genomic libraries are not representative of entire genomes. In particular, certain regions of the *H. pylori* genome appear to be toxic to *E. coli* host strains and appear to be poorly represented or unclonable. The *cagA* (Covacci *et al.*, 1993) and *vacA* (Telford *et al.*, 1994) genes were examples of this phenomenon. In addition, three regions of the *H. pylori* genome appeared as gaps in the ordered cosmid map constructed

by Bukanov and Berg (1994). Similar gaps have also been observed during the construction of ordered genomic libraries of *E. coli* (Kohara *et al.*, 1987) and *Mycobacterium leprae* (Eiglmeier *et al.*, 1993).

I have successfully constructed genomic libraries of *H. pylori* UA802 DNA using the lambda vector λ-DASH II in the host strain *E. coli* SRB. Although my results indicate that these libraries are not representative of the entire genome, they were used in the isolation of genes encoding a P-type ATPase and the rRNA of *H. pylori* UA802. To circumvent problems associated with variable representation of certain regions of the genome, the concurrent construction of libraries with other restriction enzymes or other vector systems such as Lorist6 (Bukanov and Berg, 1994) could be done. Amplification of these libraries through an alternative host strain, such as *E. coli* ER1792, may improve efficiency by reducing the effects of the restriction modification barrier.

A.2. Cloning and characterization of the *H. pylori* UA802 *hpCopA*, *hpCopP*, and rRNA genes

The P-type ATPase genes

Another of the objectives of my research was to clone the putative P-type ATPase of *H. pylori* UA802. Twenty-three clones were isolated which hybridized to a probe specific for the *H. pylori* P-type ATPase, from the λ-DASH II genomic libraries, as described in Chapter 4. A 2.7 kb fragment from

one of these clones, λP14, was subcloned in pBluescript SK$^+$ and characterized by Ge *et al.*, 1995. Three open reading frames were found by Ge *et al.*. The first ORF encoded a protein, designated HpCopA, which showed extensive homology with a copper transport protein, CopA, from *Enterococcus hirae* (Odermatt *et al.*, 1993). The second ORF encoded a protein, designated HpCopP, which had homology with the heavy-metal-binding protein MerP (Barrineau *et al.*, 1984). No homology could be determined for the protein encoded by the third ORF. Knock-out mutants of *hpCop*A and *hpCop*P were constructed by Ge *et al.* and showed increased sensitivity to copper ions, while having similar susceptibility to other heavy metals and omeprazole. The knock-out mutants were susceptible to omeprazole at levels similar to that of the wild-type, which was not expected. Since omeprazole targets the gastric human P-type ATPases, it was thought that omeprazole would also target the P-type ATPase of *H. pylori*. Data obtained by Ge *et al.* showed that the HpCopA/P proteins were not affected by omeprazole *in vitro* and that these proteins were involved in copper transport.

Ribosomal RNA genes

The cloning and characterization of the rRNA genes of *H. pylori* UA802 were also objectives of this research. A 16S gene and both copies of the 23S and 5S genes were cloned from the λ-DASH II libraries of *H. pylori* UA802 DNA. Extensive characterization of the 16S gene of *H. pylori* has already been published (Romaniuk *et al.*, 1987; Eckloff *et al.*, 1994),

so the focus of my work was the characterization of the 23S and 5S genes. Sequence analysis showed that the 23S and 5S genes were closely linked and appeared to be part of an operon, separate from the 16S gene. Three Pribnow box sequences were located in the region 5' to the 23S gene (Figure 4-10). In *E. coli*, two tandem promoters are present for each rRNA operon (Young and Steitz, 1979; de Boer *et al.*, 1979; Jinks-Robertson and Nomura, 1992). Experimental data with promoter-terminater fusion plasmids indicated that, of the two promoters, the P1 promoter was responsive to growth rate control, while transcription from the P2 promoter was constitutive. *In vivo* studies have indicated that initiation of transcription originates mostly from the P1 promoter (Sarmientos and Cashel, 1983), although *in vitro* the promoters have been observed to have relatively equal efficiency (Glaser and Cashel, 1979). This has led to speculation that the two promoters are differentially regulated and respond to different growth conditions (Jinks-Robertson and Nomura, 1992). The fact that three Pribnow boxes could be identified in the 5'-region of the 23S/5S genes of *H. pylori* UA802 suggests that similar transcriptional control may exist in *H. pylori*. Further study of the 5'-region of the *H. pylori* UA802 23S-5S operon, perhaps using promoter fusion plasmids as were used by Condon *et al.* (1992) with *E. coli*, are needed to determine how transcription is regulated, where it is initiated and what

other regions in the 5'-region are important for transcription.

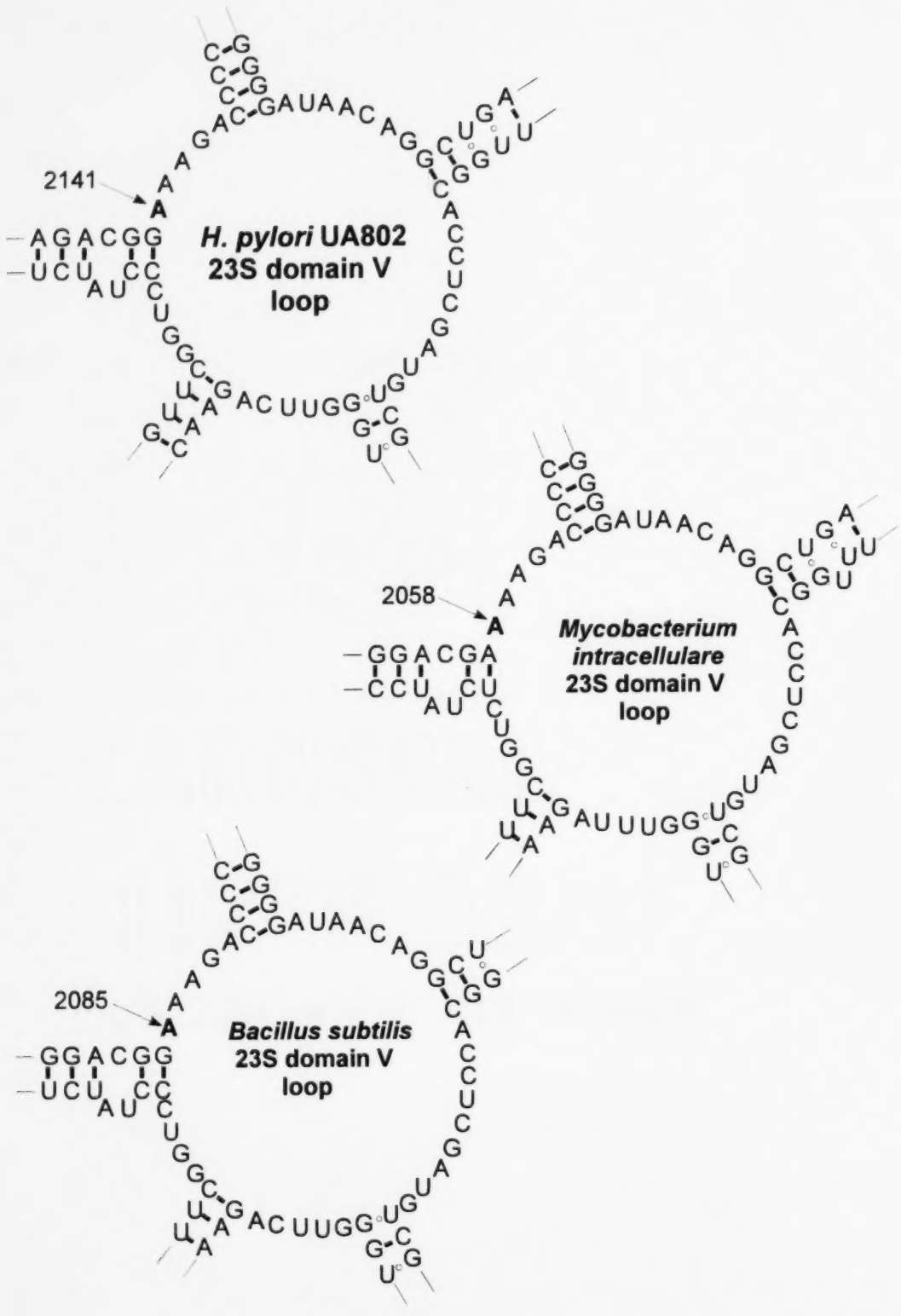
Several antibiotic resistances are specified in rRNA sequence. An example of this is the resistance to the macrolide-lincosamide-streptogramin B (MLS)-group of antibiotics such as clarithromycin and erythromycin (Skinner *et al.*, 1983; Meier *et al.*, 1994; Zhong *et al.*, 1995). In *E. coli* (Skinner *et al.*, 1983) and *B. subtilis* (Zhong *et al.*, 1995), the methylation of a specific A-residue in a conserved loop of the 23S rRNA molecule, designated domain V, has been shown to confer resistance to the MLS-group of antibiotics. The methylation is carried out by the methyltransferase ErmC', which can be acquired on a plasmid. The A-2058 residue of *E. coli* (Skinner *et al.*, 1983) and the A-2085 residue of *B.subtilis* (Zhong *et al.*, 1995) have been shown to be the methylation site of this enzyme. Recently, resistance to the MLS-group of antibiotics has been shown to arise from the mutation of the A-residue of *M. intracellulare* (equivalent to the A-2058 in *E. coli*) to G,C, or T (Meier *et al.*, 1994). The domain V loop in the *H. pylori* 23S gene sequence is almost identical to those found in *B. subtilis* and

M. intracellulare (Figure 7-1). The A-2141 residue (2141 bp from the start of the 23S gene or position 2514 on Figure 4-10) is equivalent to the A residues found in the domain V loops of the other organisms. Thus, it is possible that *H. pylori* could acquire resistance to the MLS-group of

Figure 7-1. Comparison of the domain V loops of the 23S rRNA molecules from *H. pylori* UA802, *Bacillus subtilis* and *Mycobacterium intracellulare*

Arrow indicate the A residues which can be methylated by ErmC'-methylases or mutated to confer resistance to the MLS-group of antibiotics.

The *B. subtilis* loop was from Zhong *et al.*, 1995.
The *M. intracellulare* loop was from Meier *et al.*, 1994.


2141
H. pylori UA802
23S domain V
loop
2058
Mycobacterium
intracellulare
23S domain V
loop
2085
Bacillus subtilis
23S domain V
loop

antibiotics by the acquisition of the specific methylase (ErmC') or by mutation of residue A-2141.

The complete and partial sequences of the 23S gene and the complete sequences of the 5S gene, determined for this research, were compared with 23S and 5S sequences of several other procaryotes and firmly established the phylogenetic position of *H. pylori* relative to other Eubacteria. The comparisons of the complete 23S and 5S genes confirmed that, phylogenetically, *H. pylori* and *Campylobacter* spp. were similarly distant from other Eubacteria. In addition, the comparisons of the partial 23S sequences confirmed that *H. pylori* belonged in a sub-class distinct from *Campylobacter* spp.. These findings correlated with earlier studies of the phylogeny of *Campylobacter* spp. and related organisms, including *H. pylori* (Lau *et al.*, 1987; Vandamme *et al.*, 1991; Vandamme and de Ley, 1991, Trust *et al.*, 1994).

A.3. Genomic restriction map and PFGE analyses of *H. pylori* UA802

Genomic restriction map of *H. pylori* UA802

A genomic restriction map of *H. pylori* UA802 was constructed which differed from a previously published map of Taylor *et al.* (1992a). The revised map included *Sfi*I restriction digestion data and had an additional *Nru*I fragment (*Nru*I fragment 9, Figure 5-3) and an additional *Not*I fragment (*Not*I fragment 5b, Figure 5-3). In addition, the relative positions of some of the fragments were different. The size of the genome of *H. pylori* UA802, estimated from the

PFGE data, was 1695 kb ± 17. This was in agreement with the previously estimated size of 1710 kb ± 12 for *H. pylori* UA802 determined by Taylor *et al.* (1992a), as well as with the sizes of the genomes of other *H. pylori* strains which ranged in size from 1674-1730 kb (Bukanov and Berg, 1994; Jiang, 1994).

When the *H. pylori* UA802 map was compared with genomic maps of three other *H. pylori* strains (Figure 5-6), significant variability in the arrangement of genes was observed. Even with this variability, some genes appeared to be clustered together. The *hpCop*AP genes and the *gyr*B genes appeared to cluster together in all the strains for which they were mapped, and the *vac*A, *kat*A, *hpa*A, *pfr*, and, with the exception of strain UA861, *cag*A and *fla*A genes had similar arrangements in all strains. The *vac*A, *kat*A, *hpa*A, *cag*A and and *fla*A have been implicated in the pathogenesis of *H. pylori* (Telford *et al.*, 1994; Tummuru, *et al.*, 1993; Evans *et al.*, 1993; Cover and Blaser, 1995). In a review of chromosome organization, Krawec and Riley (1990) suggest that clustering of genes may indicate that their arrangement may affect the function or regulation of the genes. This may be of particular importance for genes such as *vac*A and *cag*A in *H. pylori*, which are both variably expressed from strain to strain. Through the use of genomic restriction maps, regions which may be important for rearrangement could be localized and these regions could be isolated from genomic libraries to further dissect these regions. This may lead to an

understanding of some of the mechanisms of genomic variability, as well as variability in the mechanisms of pathogenesis of different strains of *H. pylori*.

B. Analysis of the *C.jejuni* DNA probes and MOMP

B.1. Characterization of the *C.jejuni* DNA probes

Characterization of the two *Campylobacter* diagnostic probes (Taylor and Hiratsuka, 1990) was another objective of this research. Sequencing analyses revealed that pDT1720 encoded a partial open reading frame, ORF3. No function could be determined for this partial protein sequence. Sequencing of the cloned DNA of pDT1719 showed that two partial ORFs, ORF1 and ORF2, were encoded by this plasmid. ORF1 had homology with the heat shock/serine protease genes of a number of organisms. ORF 2 exhibited homology with the *omp*R and *omp*R-like genes of the response regulators of two-component signalling systems. To my knowledge, no htr/serine protease activity has been associated with a response regulatory element as part of a two-component regulatory system. In most instances the sensory element is a protein kinase (Stock *et al.*, 1990).

The gene sequence of the *C. jejuni* MOMP has yet to be determined. The MOMP gene has not been cloned and may be an example of a toxic element (Taylor and Hiratsuka, 1990). As mentioned earlier, the involvement of the MOMP porin in the virulence of *C. jejuni* is not clear, although it has been

implicated in adhesion to eucaryotic cells (de Melo and Pechere, 1991).

The N-terminal sequence of the first 30 amino acids of the purified MOMP was determined. This sequence had significant homology with the *C. rectus* 43 kb and 51 kb major outer membrane proteins. *C. rectus* was formerly called *Wolinella recta*. This organism was recently reclassified and found to belong to *Campylobacter* spp. (Vandamme *et al.*, 1991). The comparisons of the N-terminal amino acid sequence of the *C. jejuni* MOMP to the *C. rectus* MOMPs (Figure 6-5), supported the inclusion of *C. rectus* in the family *Campylobacteriaceae* (Vandamme *et al.*, 1991). In addition, because the *C. rectus* was the only sequence in the database which showed such a degree of homology with the *C. jejuni* MOMP, the N-terminal amino acid sequence of the MOMP may be a possible candidate for a species-specific sequence which can be used for classification of *Campylobacter* species.

C. Future study

There are several possibilities for further research stemming from the results presented in this thesis. The genomic libraries of *H. pylori* UA802 DNA that I constructed using the vector λ-DASH will be useful as a source of cloned DNA for future work.

The study of the transcriptional control mechanisms of the rRNA genes of *H. pylori* would be interesting. As mentioned earlier, the rRNA genes of most procaryotes are co-

transcribed from a single operon and this large transcript is then processed to give the individual species of rRNA. In this way, the relative amount of each rRNA species is kept equal. In *H. pylori*, the 23/5S genes and the 16S genes are transcribed separately. Thus, the amount of each rRNA species must be regulated by the transcriptional control mechanism. Cloning and characterization of the 5'-flanking region of the 23S/5S operons and the 16S genes would help in determining how transcription of these rRNA genes is regulated in *H. pylori*.

Finally, sequencing of the *htr*AB genes of *C. jejuni* UA580 could be completed. The clone pDT1719 could be used to screen genomic libraries of *C. jejuni* UA580 to obtain the complete genes. Since the HtrA protein showed significant homology to the HtrA heat shock proteins from *E. coli* and *S. typhimurium*, it is possible that this protein responds to heat shock. Does this element respond directly to heat shock or is it part of a cascade of reactions responding to heat shock? Transcription of the *htr*A gene of *E. coli* is controlled by a novel sigma factor, σ^E (Erickson and Gross, 1989), which is expressed in response to heat shock. It would be interesting to determine if the *C. jejuni htrA* gene is under similar control or if it responds directly to some environmental stimulus.

If these two proteins are involved in a two-component regulatory system, it would also be interesting to determine how the signal is transduced from the sensory element (*htr*A)

to the response element (*htr*B). The *C. jejuni* HtrA protein exhibits homology with serine proteases, so it may activate the response element by proteolytic cleavage. It would be interesting to determine what functions are controlled by the response element once it is activated. The complete sequences of these genes may provide more information about the function of these proteins.

The N-terminal sequence of the *C. jejuni* UA580 MOMP can be used to design oligonucleotide probes to screen DNA libraries of *C. jejuni* for the isolation of a gene encoding this protein. An alternative to this would be to use the oligonucleotide as a PCR primer, using random primers in the opposite direction. Cloning and sequencing of any products of this PCR would reveal if sections of the MOMP gene were cloned, and these cloned PCR products could then be used to probe *C. jejuni* DNA libraries to isolate the remainder of the gene. Part or all of the cloned gene of the MOMP may be useful as a genus-specific probe for *Campylobacter* species, similar to the *C. jejuni* DNA probe described by Strucki *et al.* (1995). Strucki *et al.* described the cloning of a *C. jejuni* membrane protein gene, which encoded a 24 kDa protein. This gene, when used as a DNA probe, could be used to distinguish *C. jejuni* from other *Campylobacter* species.

Bibliography

Akopyanz, N., N.O. Bukanov, T.U. Westblom, and D.E. Berg. 1992a. PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen *Helicobacter pylori*. Nucleic Acids Res **20**:6221-6225.

Akopyanz, N., N.O. Bukanov, T.U. Westblom, S. Kresovich, and D.E. Berg. 1992b. DNA diversity among clinical isolates of *H. pylori* detected by PCR-based RAPD fingerprinting. Nucleic Acids Res **20**:5137-5142.

Akopyants, N., K.A. Eaton, and D.E. Berg. 1995. Adaptive mutation and cocolonization during *Helicobacter pylori* infection of gnotobiotic piglets. Infect Immun **63**(1):116-121.

Altshul, S.F., W. Gish, W. Miller, E.W. Meyers, and D.J. Lipman. 1990. Basic local alignment search tool. J Mol Biol **215**:403-410.

Ausubel, F.M., R. Brent, R.E. Kingston, P.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1987. Introduction of plasmid DNA into cells. *In*: Current protocols in molecular biology, Wiley Interscience, New York, NY, USA. pp 1.6.2-1.6.4.

Barrineau, P., P. Gilbert, W.J. Jackson, C.S. Jones, A.O. Summers, and S. Wisdom. 1984. The DNA sequence of the mercury resistance operon of the IncFII plasmid NR1. J Mol Appl Genet **2**:601-619.

Baxter, A., C.J. Campbell, D.M. Cox, C.J. Ginham, and J.E. Pendlebury. 1989. Proteolytic activities of human *Campylobacter pylori* and ferret gastric *Campylobacter*-like organism. Biochem Biophys Res Com **163**:1-7.

Béji, A., F. Mégraud, P. Vincent, F. Gavini, D. Izard, and H. Leclerc. 1988. GC content of DNA of *Campylobacter pylori* and other species belonging or related to the genus *Campylobacter*. Ann Inst Pasteur/Microbiol **139**:527-534.

Birnboim, H.C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res **7**:1513-1523.

Blaser, M. and L.B. Reller. 1981. *Campylobacter* enteritis. New Engl J Med **305**(24):1444-1452.

Blaser, M.J. 1987. Gastric *Campylobacter*-like organism, gastritis and peptic ulcer disease. Gastroenterology **93**:371-383.

Blaser, M.J. 1992. Hypothesis on the pathogenesis and natural history of *Helicobacter pylori* induced inflammation. Gastroenterology **102**:720-727.

Borén, T., P. Falk, K.A. Roth, G. Larson, and S. Normark. 1993. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. Science **262**:1892-1895.

Brosius, J., T.J. Dull, and H.F. Noller. 1980. Complete nucleotide sequence of a 23S ribosomal RNA gene from *Escherichia coli*. Prc Natl Acad Sci USA **77**:201-204.

Brown.ee, G.G., F. Sanger, and B.G. Barrell. 1967. Nucleotide sequence of 5S-ribosomal RNA from *Escherichia coli*. Nature (London) **215**:735-736.

Buck,G.E. 1990. *Campylobacter pylori* and gastroduodenal disease. Clin Microbiol Rev **3**:1-12.

Bugnoli, M., P.F. Bayeli, R. Rappuoli, C. Pennatini, N. Figura, and J.E. Crabtree. 1993. Inhibition of *Helicobacter pylori* urease by omeprazole. Eur J Gastroenterol Hep **5**:683-685.

Bukanov, N.O., and D.E. Berg. 1993. Ordered Cosmid Library and high resolution physical genetic map of *Helicobacter pylori* strain NCTC 11638. Mol Microbiol **11(3)**:509-523.

Chauthaiwale, V.M., A. Therwath, and V.V. Deshpande. 1992 Bacteriophage lambda as a cloning vector. Microbiol Rev **56**(4):577-591.

Chen, J. and D.A. Morrison. 1987. Cloning of *Streptococcus pneumoniae* DNA fragments in *Escherichia coli* requires vectors protected by strong transcriptional terminators. Gene **55**:179-187.

Chen, X.G., P. Correa, J. Offerhaus, E. Rodriguez, F. Janney, E. Hoffmann, J. Fox, F. Hunter and S. Diavolitsis. 1986. Ultrastructure of the gastric mucosa harboring *Campylobacter*-like organisms. Am J Clin Pathol **86**:575-582.

Chen, X., and L.R. Finch. 1989. Novel arrangement of rRNA genes in *Mycoplasma gallisepticum*: separation of the 16S gene of one set from the 23S and 5S genes. J Bacteriol **171**:2876-2878.

Clayton, C.L., B.W. Wren, P. Mullany, A. Topping, and S. Tabaqchali. 1989a. Molecular cloning and expression of *Campylobacter pylori* species-specific antigens in *Escherichia coli* K-12. Infect Immun **57(2)**:623-629.

Clayton, C.L., M.J. Pallen, H. Kleanthous, B.W. Wren, and S. Tabaqchali. 1989b. Nucleotide sequence of two genes from *Helicobacter pylori* encoding for urease subunits. Nucleic Acids Res **18(2)**:362.

Condon, C., J. Philips, Z. Fu, C. Squires, and C.L. Squires. 1992. Comparison of the expression of the seven ribosomal RNA operons in *Escherichia coli*. EMBO J **11**(11):4175-4185.

Correa, P., J. Fox, E. Fontham, B. Riuz, Y.P. Lin, D. Zavala, N. Taylor, D. Mackinley, E. deLima, H. Portilla, and G. Zarama. 1990. *Helicobacter pylori* and gastric carcinoma. Serum antibody prevalence in populations with contrasting cancer risks. Cancer **66**:2569-2574.

Corry, M.J. P.I. Payne, and T.A. Dyer. 1974. The nucleotide sequence of 5S rRNA from the blue-green alga *Anacystis nidulans*. FEBS leet **46**(1):63-66.

Covacci, A., S., Censini, M Bugnoli, R. Petracca, D. Burroni, G. Macchia, A. Massone, E. Papini, Z. Xiang, N. Figura, and R. Rappuoli. 1993. Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer. Proc Natl Acad Sci USA **90**:5791-5795.

Cover, T.L., C.P. Dooley, and M.J. Blaser. 1990. Characterization of and human serologic response to proteins in *Helicobacter pylori* broth culture supernatants with vacuolizing cytotoxin activity. Infect Immun **58**:603-610.

Cover, T.L. and M.J. Blaser. 1992a. *Helicobacter pylori* and gastroduodenal disease. Ann Rev Med **43**:135-145.

Cover, T.L. and M.J. Blaser. 1992b. Purification and characterization of the vacuolating toxin from *Helicobacter pylori*. J Biol Chem **267**:10570-10575.

Cover, T.L., M.K.R Tummuru, P. Cao, S.A. Thompson, and M.J. Blaser. 1994. Divergence of genetic sequences for the vacuolating toxin among *Helicobacter pylori* strains. J Biol Chem **269(14)**:10566-10573.

Cover, T.L., and M.J. Blaser. 1995. *Helicobacter pylori*: A bacterial cause of gastritis, peptic ulcer disease, and gastric cancer. ASM News **61**(1):21-26.

Crabtree, J.E., J.D. Taylor, J.I. Wyatt, R.V. Heatley, T.M. Shallcross, D.S. Tomkins, and B.J. Rathbone. 1991. Mucosal IgA recognition of *Helicobacter pylori* 120 kDa protein, peptic ulceration, and gastric pathology. Lancet **338**:332.

Crabtree, J.E., T.M. Shallcross, R.V. Heatley and J.I. Wyatt. 1992. Expression of 120-kilodalton protein and cytotoxicity in *Helicobacter pylori*. J Clin Pathol **45**:733-736.

Crabtree, J.E., S.M. Farmery, I.J.D. Lindley, N. Figura, P. Peichl, and D.M. Tomkins. 1994. CagA/cytotoxic strains of *Helicobacter pylori* and interleukin-8 in gastric epithelial cell lines. J Clin Pathol **47**:945-950.

Cussac, V., R.L. Ferrero, and A. Labigne. 1992. Expression of *Helicobacter pylori* urease genes in *Escherichia coli* grown under nitrogen-limiting conditions. J Bacteriol **174**:2466-2473.

de Boer,H.A., S.F. Gilbert, and M. Nomura. 1979. DNA sequences of promoter regions of rRNA operons *rrnE* And *rrnA* in *E. coli*. Cell **17**:201-209.

de Melo, M.A. and J. Pechere. 1990. Identification of *Campylobacter jejuni* surface proteins that bind eucaryotic cells *in vitro*. Infect Immun **58**(6):1749-1756.

Dick, J.D. 1990. *Helicobacter* (*Campylobacter*) *pylori*: A new twist to an old disease. Ann Rev Microbiol **44**:249-269.

Dixon, M.F. 1991. *Helicobacter pylori* and peptic ulceration: histopathological aspects. J Gastroenterol Hep **6**:125-130.

Douglas, S.E. and W.F. Doolittle. 1984. Complete nucleotide sequence of the 23S ribosomal RNA gene of the cyanobacterium *Anacystis nidulans*. Nucleic Acids Res **12**(7):3373-3386.

Dytoc, M., B. Gold, M. Louie, M Huesca, L. Fedorko, S. Crowe, C. Lingwood, J. Brunton, and P. Sherman. 1993. Comparison of *Helicobacter pylori* and attaching-

effacing *Escherichia coli* adhesion to eukaryotic cells. Infect Immun **61**(2):448-456.

Eaton, K.A., D. R. Morgan, and S. Krakowka. 1989. *Campylobacter pyloridis* virulence factors in gnotobiotic piglets. Infect Immun **57**:1119-1125.

Eaton, K.A., C.L. Brooks D. R. Morgan, and S. Krakowka. 1991. Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. Infect Immun **59**:2470-2475.

Eckloff, B.W., R.P. Podzorski, B.C. Kline, and F.R Cockerill III. 1994. A comparison of 16S ribosomal DNA sequences from five isolates of *Helicobacter pylori*. Int J Syst Bacteriol **44(2)**:320-323.

Egebjerg, J., N. Larsen, and R.A. Garrett. 1990. Structural map of 23S rRNA. *In*: W.E. Hill, A. Dahlberg, R.A. Garrett, P.B. Moore, D. Schlessinger, and J.R. Warner (eds). The Ribosome: Structure, function and evolution. American Society of Microbiologists, Washington, D.C. USA. pp.168-179.

Eiglmeier, K. N. Honoré, S.A. Woods, B. Caudron, and S.T. Cole. 1993. Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. Mol Microbiol **7**(2):197-206.

Ellis, R.E., J.E. Sulston, and A.R. Coulson. 1986. The r·DNA of *Caenorhabditis elegans*: sequence and structure. Nucleic Acids Res **14**:2345-2364.

Erdmann, V.A., J. Wolters, E. Huysmans, and R. de Wachter. 1985. Collection of published 5S, 5.8S and 4.5S ribosomal RNA sequences. Nucleic Acids Res **13**:r105-r153.

Erickson, J.W., and C.A. Gross. 1989. Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternate σ factor involved in high temperature gene expression. Gen Dev. **3**:1462-1471.

Evans, D.G., D.J. Evans, Jr., J.J. Moulds, and D.Y. Graham. 1988. *N*-acetylneuraminyllactose-binding fibrillar hemagglutinin of *Campylobacter pylori*: A putative colonization factor antigen. Infect Immun **56(11)**:2896-2906.

Evans, D.J., Jr., D.G. Evans, K.E. Smith, and D.Y. Graham. 1989a. Serum antibody responses to the *N*-

acetylneuraminyllactose-binding hemagglutinin of *Campylobacter pylori*. Infect Immun **57(3)**:664-667.

Evans, D.G., D.J. Evans, Jr., and D.Y. Graham. 1989b. Receptor-mediated adherence of *Campylobacter pylori* to mouse Y-1 adrenal cell monolayers. Infect Immun **57(8)**:2272-2278.

Evans, D.G., T.K. Karjalainen, D.J. Evans, Jr., D.Y. Graham, and C Lee. 1993. Cloning, nucleotide sequence, and expression of a gene encoding an adhesin subunit protein of *Helicobacter pylori*. J Bacteriol **175(3)**:674-683.

Ezaki, T., N. Takeuchi, S Liu, A, Kai, H. Yamamoto, and E. Yabuuchi. 1988. Small-scale DNA preparation for rapid genetic identification of *Campylobacter* species without radioisotope. Microbiol Immun **32**:141-150.

Fellinius, E., T. Berglindh, G. Sachs, L. Olbe, B. Elander, S.E. Sjostrand, and B. Wallmark. 1981. Substituted benzimidazoles inhibit gastric acid secretion blocking (H^++K^+)-ATPase. Nature (London) **290**:159-160.

Feng,D.f., and R.F. Doolittle. 1987. Progressive sequence alignment as a prerequisite to create phylogenetic trees. J Mol Evol **25**(4):351-360.

Ferrero,R.L., V. Cussac, P. Courcoux, and A. Labigne. 1992. Construction of isogenic urease-negative mutants of *Helicobacter pylori* by allelic exchange. J Bacteriol **174**:4212-4217.

Figura, N., P. Guglielmetti, A. Rossolini, A. Barberi, G. Cusi, R.A. Musmanno, M. Russi, and S. Quaranta. 1989. Cytotoxin production by *Campylobacter pylori* strains isolated from patients with peptic ulcers and from patients with chronic gastritis only. J Clin Microbiol **27**:225-226.

Forman, D., F. Sitas, D.G. Newell, A.R. Stacey, J. Boreham, R. Peto, T.C. Campbell, J. Li, and J. Chen. 1990. Geographic association between infection of *Helicobacter pylori* antibody prevalence and gastric cancer mortality in rural China. Int J Cancer **46**:608-611.

Fonstein, M. and R. Haselkorn. 1995. Physical mapping of bacterial genomes. J Bacteriol **177**(12):3361-3369.

Forman, D., D.G. Newell, F. Fullerton, J.W. Yarnell, A.R. Stacey, and F. Sitas. 1991. Association between infection with *Helicobacter pylori* and risk of

gastric cancer: evidence from a prospective investigation. Brit Med J. **302**:1302-1305.

Foxall, P.A., L. Hu and H.L.T. Mobley. 1992. Use of polymerase chain reaction-amplified *Helicobacter pylori* urease structural genes for differentiation of isolates. J Clin Microbiol **30**:739-741.

Frazier, B.A., J.D. Pfeifer, D.G. Russell, P. Falk, A.N. Olsén, M. Hammar, T.U. Westblom, and S.J. Normark. 1993. Paracrystalline inclusions of a novel ferritin containing nonheme iron, produced by the human gastric pathogen *Helicobacter pylori*: Evidence for a third class of ferritins. J Bacteriol **175(4)**:966-972.

Fukunaga, M., and I. Mifuchi. 1989. Unique organization of *Leptospira interrogans* rRNA genes. J Bacteriol **171**:5763-5767.

Fukunaga, M., I. Horie, and I. Mifuchi. 1990. Isolation and characterization of the 5S rRNA gene in *Leptospira interrogans*. J Bacteriol **172**:3264-3268.

Fujimoto, S., B. Marshall, and M.J. Blaser. 1994. PCR-based restriction fragment length polymorphism typing of *Helicobacter pylori*. J Clin Microbiol **32**:331-334.

Gazumayan, A., J.J. Schwartz, D. Liveris, and L. Schwartz. 1994. Sequence analysis of the ribosomal RNA operon of the lyme disease spirochete, *Borrelia burgdorferi*. Gene **146**(1):57-65.

Ge, Z., K. Hiratsuka, and D.E. Taylor. 1995. Nucleotide sequence and mutational analysis indicate that two *Helicobacter pylori* genes encode a P-type ATPase and a cation-binding protein associated with copper transport. Mol Microbiol **14**:97-106.

Gellini, L., L. Marzio, A.D. Girolamo, N. Allocati, L. Grossi, and B. Dainelli. 1991. Enhanced clearing of *Helicobacter pylori* after omeprazole plus roxithromycin treatment. FEMS Microbiol Lett **84**:235-258.

Gibson,T.J., A.R. Coulson, J.E. Sulston and P.F.R. Little. 1987. Lorist2, a cosmid with transcriptional terminators insulating vector genes from interference by promoters within the insert: effect on DNA yield and cloned insert frequency. Gene **53**:275-281.

Gilbert, S.F., H. A. de Boer, and M. Nomura. 1979. Identification of initiation sites for the *in Vitro*

transcription of rRNA operons *rrnE* and *rrnA* in *E. coli*. Cell **17**:211-224.

Glaser, G. and M. Chashel. 1979. *In vitro* transcripts from the *rrn*B ribosomal RNA cistron originate from two tandem promoters. Cell **16**:111-121.

Goodwin, C.S., R.K., McCulloch, J.A. Armstrong, and S.H. Wee. 1985. Unusual cellular fatty acids and distinctive ultrastructure in the new spiral bacterium (*Campylobacter pyloridis*) from the human gastric mucosa. J Med Microbiol **9**:257-267.

Goodwin, C.S., J.A. Armstrong, and B.J. Marshall. 1986. *Campylobacter pyloridis*, gastritis and peptic ulceration. J Clin Pathol **39**:353-365.

Goodwin, C.S., J.A. Armstrong, T. Chivers, M. Peters, M.D. Colins, L. Sly, W. McConnell, and W.E.S. Harper. 1989. Transfer of *Campylobacter pyloridis* and *Campylobacter mustelae* to *Helicobacter* gen. nov. as *Helicobacter pylori* comb. nov. and *Helicobacter mustelae* comb. nov. respectively. Int J Syst Bacteriol **39**:397-405.

Goodwin, C.S., and J.A. Armstrong. 1990. Microbiological Aspects of *Helicobacter pylori* (*Campylobacter pylori*). Eur J Clin Microbiol Infect Dis **9(1)**:1-13.

Goodwin, C.S. and B.W. Worsley. 1993. The *Helicobacter* Genus: History of *H. pylori* and taxonomy of current species. *In*: C.S. Goodwin and B.W. Worsley (eds). *Helicobacter pylori*: Biological and Clinical Practice. CRC Press, Inc., USA. pp.1-14.

Gough, J.A., and N.E. Murray. 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J Mol Biol **166**:1-19.

Green, C.J. G.C. Stewart, M.A. Hollis, B.S. Vold and K.F. Bott. 1985. Nucleotide sequence of the *Bacillus subtilis* ribosomal RNA operon *rrn*B. Gene **37**:261-266.

Guenzi, E., A.M. Gac, M.A. Sicard and R. Hakenbeck. 1994. A two-component signal-transducing system involved in competence and penicillin susceptibility in laboratory mutants of *Streptococcus pneumoniae*. Mol Microbiol **12**(3):505-515.

Gutell, R.R., M.N. Schnare, and M.W. Gray. 1992. A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures. Nucleic Acids Res **19**:1156

Haas, R., T.F. Meyer, and J.P.M. van Putten. 1993. Aflagellated mutants of *Helicobacter pylori* generated by genetic transformation of naturally competent strains using transposon shuttle mutagenesis. Mol Microbiol **8**:753-760.

Hagblom, P., E. Segal, E. Billyard, and M. So. 1985. Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. Nature (London) **315**(6015):156-158.

Hartmann, R.K., and V.A. Erdmann. 1989. *Thermus thermophilus* 16S rRNA is transcribed from an isolated transcriptional unit. J Bacteriol **171**:2933-2941.

Hazell, S.L. and A. Lee. 1986. *Campylobacter pyloridis*, urease, hydrogen ion back diffusion, and gastric ulcers. Lancet **ii**:15-17.

Hazell, S.L., A. Lee, L. Brady, and W. Hennessy. 1986. *Campylobacter pyloridis* and gastritis: association with intercellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. J Infect Dis **153**:658-663.

Hebbel, J. 1943. Chronic gastritis: Its relation to gastric and duodenal ulcer and to gastric cancer. Am J Pathol **19**:43-71.

Hemalatha, S.G., B. Drumm, and P. Sherman. 1991. Adherance of *Helicobacter pylori* to human gastric epithelial cells *in vitro*. J Med Microbiol **35**:197-202.

Hesse, J.E., L. Wieczorek, K. Altendorf, A.S. Reicin, E. Dorus and W. Epstein. 1984. Sequence homology between two membrane transport ATPases, the Kdp-ATPase of *Escherichia coli* and the Ca^{2+}-ATPase of sarcoplasmic reticulum. Proc Natl Acad Sci USA **81**:4746-4750.

Hiratsuka, K. 1987. Sequencing of *Azotobacter vinelandii nif* DNA. MSc thesis. University of Alberta, Edmonton, Canada

Hopfl, P., W. Ludwig, and K.H. Schleifer 1988. Complete nucleotide sequence of a 23S ribosomal RNA gene from *Rhodobacter capsulatus*. Nucleic Acids Res. **16**(5):2343.

Hopfl, P., W. Ludwig, K.H. Schleifer, and M. Larsen. 1989. The 23S ribosomal RNA high-order structure of *Pseudomonas cepacia* and other prokaryotes. Eur J Biochem **185**(2):355-364.

Hori, H., S. Osawa, K. Murao, and H. Ishikura. 1980. The nucleotide sequence of 5S ribosomal RNA from *Micrococcus lysodeikticus*. Nucleic Acids Res **8**:5423-5426.

Huang, W.M. 1991. *IN*: Molecular biology of DNA topoisomerase and its application to chemotherapy. T. Andoh, H. Ikeda and M. Oguro eds. CRC Press.

Huyer, M., T.R. Parr, R.E.W. Hancock, and W.J. Page. 1986. Outer membrane porin protein of *Campylobacter jejuni*. FEMS Microbiol Lett **37**:247-250.

Iwahi, T., H. Satoh, M. Nakao, T. Iwasaki, T. Yamazaki, K. Kubo, T. Tamura, and A. Imada. 1991. Lansoprazole, a novel benzimidazole proton pump inhibitor, and its related compounds have selective activity against *Helicobacter pylori*. Antimicrobial Agents Chemother **25**:490-496.

Jiang, Q. 1994. Genomic analysis of *Campylobacter* and *Helicobacter*. MSc thesis. University of Alberta, Edmonton, Canada.

Kato, S. and K. Komagata. 1986. Nucleotide sequence of 5S ribosomal RNAs of *Protomonas extorquens*, *Rhodopseudomonas palustris*, *Rhodobacter capsulatus*, and *Erythrobacter longus*. Nucleic Acids Res **14**(10):4371.

Kennel, W.L. and S.C. Holt. 1991. Extraction, purification , and characterization of the major outer membrane proteins from *Wolinella recta* ATCC 33238. Infect Immun **59**(10):3740-3749.

Kessler, U., U. Maid, and K. Zetsche. 1992. An equivalent to bacterial *omp*R genes is encoded on the plastid genome of red algae. Plant Molec Biol **18**(4):777-780.

Kim. N.W., R.R. Gutell, and V.L. Chan. 1993. Complete nucleotide sequence and organization of *rrn*A ribosomal RNA operon in *Campylobacter jejuni* TGH9011 (ATCC43431). Unpublished direct Genbank submission. Accession number Z29326.

Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495-508.

Komiya, H., M. Kawakami, S. Takemura, I. Kumagai, and V.A. Erdmann. 1983 Terminal heterogeneity and corrections of the nu;cleotide sequence of 5S rRNA from

an extreme thermophile, *Thermus thermophilus* HB8. Nuclic Acids Res **11**(5):913-916.

Konkel, M.E., R.T. Marconi, D.J. Mead, and W. Cieplak, Jr. 1994. Identification and characterization of an intervening sequence within the 23S ribosomal RNA genes of *Campylobacter jejuni*. Mol Microbiol **14**(2):235-241.

Kostrzynska, M., J.D. Betts, J.W. Austin, and T.J. Trust. 1991 Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella. J Bacteriol **173(3)**:937-946.

Kostrzynska, M., P.W. O'Toole, D.E. Taylor, and T.J. Trust. 1994 Molecular characterization of a conserved 20-kilodalton membrane-associated lipoprotein antigen of *Helicobacter pylori*. J Bacteriol **176**:5938-5948.

Krawec, S. and M. Riley. 1990. Organization of the bacterial chromosome. Microbiol Rev **54**(4):502-539.

Labigne, A, V. Cussac, and P. Courcoux. 1991. Shuttle cloning and nucleotide sequences of *Helicobacter pylori* genes responsible for urease activity. J Bacteriol **173(6)**:1920-1931.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680-685.

Lambert,MN.A., C.M. Patton, T.J. Barrett, and C.W. Moss. 1987. Differentiation of *Campylobacter* and *Campylobacter*-like organisms by cellular fatty acid compositions. J Clin Microbiol **25**:2652-2655.

Lau, P.P., B. DeBrunner-Vossbrinck, B. Dunn, K. Miotto, M.T. MacDonell, D. M. Rollins, C.J. Pillidge, R. B. Hespell, R.R. Colwell, M. L. Sogin, and G.E. Fox. 1987. Phylogenetic diversity and position of the genus *Campylobacter*. Syst Appl Microbiol **9**:231-238.

Lee, A., J. Fox, and S. Hazell. 1993. Pathogenicity of *Helicobacter pylori*: A perspective. Infect Immun **61**:1601-1610.

Lee,J.W., and F.M. Hullet. 1992. Nucleotide sequence of the *pho*P gene encoding PhoP, the response regulator of the phosphate regulon of the *Bacillus subtilis*. Nucleic Acids Res **20**(21):5848.

Leffers, H., J. Kjems, L. Østergaard, N. Larsen and R.A. Garrett. 1987. Evolutionary relationships among archaebacteria. A comparative study of 23S rRNAs of a sulphur-dependent thermophile, and extreme halophile and a thermophilic methanogen. J Mol Biol **195**:43-61.

Leunk,R.D., P.T. Johnson, B.C. David,W.G. Kraft, and D.R. Morgan. 1988. Cytotoxic activity in broth culture filtrates of *Campylobacter pylori*. J Med Microbiol **26**:414-417.

Leying, H., S. Suerbaum, G. Geis and R. Haas. 1992. Cloning an genetic characterization of a *Helicobacter pylori* flagellin gene. Mol Microbiol **6(19)**:2863-2874.

Liesack, W., and E. Stackebrandt. 1989. Evidence for unlinked *rrn* operons in the planctomycete *Pirellula marina*. J Bacteriol **171**:5025-5030.

Lindberg, P., P. Nordberg, T. Alminger, A. Brandstrom, and B. Wallmark. 1986. The mechanism of action of the gastric acid secretion inhibitor omeprazole. J Med Chem **29**:1327-1329.

Lingwood, C.A. 1993. *Helicobacter pylori*: Receptors and adhesins. *In*: C.S. Goodwin and B.W. Worsley (eds). *Helicobacter pylori*: Biology and Clinical Practice. CRC Press, Inc., USA. pp.209-222.

Lior, H, and W.M. Johnson. 1985. Catalase, peroxidase and superoxide dismutase activities in *Campylobacter* spp. *In*:Pearson, A.D., M.B. Skirrow, H. Lior and B. Rowe (ed): *Campylobacter* III: Proceedings of the third international workshop on *Campylobacter* infections. Public Health Laboratory Service, London. pp.226-227.

Lipinska, B., M. Zykucz, and C. Georgopoulos. 1988. Sequence analysis and regulation of the *htr*A gene of *Escherichia coli*: evidence for a σ^{32}-independent mechanism at elevated temperatures. Nucleic Acids Res **16**:10053-10067.

Lipinska, B., O. Fayet, L. Baird, and C. Georgopoulos. 1989. Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J Bacteriol **171**(3):1574-1584.

Lipinska, B., M. Zykucz, and C. Georgopoulos. 1990. The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. J Bacteriol **172**(4):1791-1797.

Loo, V.G., P. Sherman, and A.G. Matlow. 1992 *Helicobacter pylori* infection in a pediatric population: *In vitro* susceptibilities to omeprazole and eight antimicrobial agents. Antimicrobial Agents Chemother **36**:1133-1135.

Majewski, S.I.H., and C.S. Goodwin. 1988. Restriction endonuclease analysis of the genome of *Campylobacter pyloridis* with a rapid extraction method: evidence for considerable genomic variation. J Infect Dis **157**:465-471.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *In*:Molecular Cloning: A laboratory manual. Maniatis, T., E.F. Fritsch, and J. Sambrook eds.Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marchetti, M., B. Arico, D. Burroni, N. Figura, R. Rappuoli, and P. Ghiara. 1995. Development of a mouse model of *Helicobacter pylori* infection that mimics human disease. Science **267**:1655-1658.

Marmur, J., and P. Doty. 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J Mol Biol **5**:109-118.

Marotta, C.A., F. Varricchio, I. Smith, S.M. Weissman,M.L. Sogin and N.R. Pace. 1976. The primary structure of the *Bacillus subtilis* and *Bacillus stearothermophilus* 5S ribonucleic acids: Sequence variations between polynucleotides derived from mesophilic and thermophilic organisms. J Biol Chem **251**:3122-3127.

Marshall, B.J. 1983. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet **i**:1273-1275.

Marshall, B.J., and J.R. Warren. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet **1**:1311-1314.

Marshall, B.J., H. Royce, D.I. Annear, J.W. Pearman, J.R. Warren and J.A. Armstrong. 1984. Original isolation of *Campylobacter pyloridis* from human gastric mucosa. Microbios Lett **25**:83-88.

Marshall, B.J., D.B. McGechie, P.A. Rogers, and R.J. Glancy. 1985. Pyloric *Campylobacter* infection and gastroduodenal disease. Med J Australia **142**:439-444.

Marshall, B.J., and C.S. Goodwin. 1987. Revised nomenclature of *Campylobacter pyloridis*. Int J Syst Bacteriol **37**:68.

Marshall, B.J. and I. Surveyor. 1988. Carbon-14 urea breath test for the diagnosis of *Campylobacter pylori* associated with gastritis. J Nuclear Med **29**:11-16.

Marshall, B.J.,L.J. Barrett, C. Prakash, R.W. McCallum, and R.L. Guerrant. 1990. Urea protects *Helicobacter (Campylobacter) pylori* from the bactericidal effect of acid. Gastroenterology **99**:697-702.

Matsudaira, P. 1987. Sequencing of picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J Biol Chem **262**:10035-10038.

McNulty, C.A.M. and R. Wise. 1985. Rapid diagnosis of *Campylobacter*-associated gastritis. Lancet **1(8443)**:1443-1444.

Mégraud, F., V. Neman-Simha, and D. Brügmann. 1992. Further evidence of the toxic effect of ammonia produced by *Helicobacter pylori* urease on human epithelial cells. Infect Immun **60**:1858-1863.

Meier, A. P. Kirschner, B. Springer, V.A. Steingrube, B.A. Brown, R. J. Wallace, Jr., and E.C. Böttger. 1994. Identification of mutations in 23S rRNA gene of clarithromycin-resistnt *Mycobacterium intracellulare*. Antimicrob Agents Chemother **38**()2):381-384.

Miura, K., J. Kakuchi, E. Endo, T. Ueda, K Kobayashi, M Nakao, and M. Ishimoto. 1986. Nucleotide sequence of 5S ribonucleic acid from a sulfate-reducing bacterium, *Desulfovibrio vulgaris*. Chem Pharm Bull **34**:4190-4194

Mobley, H.L.T., M.J. Cortesia, L.E. Rosenthal, and B.D. Jones. 1988. Characterization of urease from *Campylobacter pylori*. J Clin Microbiol **26**:831-836.

Moon, H.W, S.C. Whipp, R.A. Argenzio, M.M. Levine, and R. A. Giannella. 1988. Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. Infect Immun **41**:1340-1351.

Moore, R.A., B. Beckthold, S. Wong, A. Kureishi, and L.E. Bryan. 1995. Nucleotide sequence of the *gyr*A gene and characterization of ciprofloxacin-resistant mutants

of *Helicobacter pylori*. Antimicrobial Agents Chemother **39**(1):107-111.

Morrison, D.A. and B. Jaurin. 1990. *Streptococcus pneumoniae* possesses canonical *Escherichia coli* (sigma 70) promotors. Mol Microbiol **4**(7):1143-1152.

Nagasawa,S., K. Ishige, and T. Mizuno. 1993. Novel members of the two-component signal transduction gene in *Escherichia coli*. J Biochem **114**:350-357.

Neidhardt, R.C., and R. VanBogelen. 1987. Heat shock response. *In*: F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (eds): *Escherichia coli* and *Salmonella typhimurium*. Cellular and Molecular Biology. pp 1134-1345.

Newell, D.G., P. Nuijten, A.R. Stacey, and S.L. Hazell. 1991. The cloning, sequencing and antigenicity of *Helicobacter pylori* catalase. Ital J Gastroentorol **23**(suppl. 2):7.

Newnham, P.J., and D.E. Taylor. 1994. Molecular analysis of RepHI1A, a minimal replicon of the IncHI1 plasmid R27. Mol Microbiol **11**(4):757-768.

Nomura, M., R. Gourse and G. Baughman. 1984. Regulation of the synthesis of ribosomes and ribosomal components. Ann Rev Biochem **53**:75-117.

Odermatt,A., H. Suter, R. Krapf, and M. Solioz. 1993. Primary structure of two P-type ATPases involved in copper homeostasis in *Enterococcus hirae*. J Biol Chem **268**:12775-12779.

Ogasawara, N., S. Nakai, and H. Yoshikawa. 1994a. Systematic sequencing of the 180 kb region of the *Bacillus subtilis* chromosome containing the replication origin. DNA Res **1**:1-14.

O'Toole, P.W., S.M Logan, M. Kostrzynska, T. Wadström, and T.J. Trust. 1991. Isolation and biochemical and molecular analyses of a species-specific protein antigen from the gastric pathogen *Helicobacter pylori*. J Bacteriol **173(2)**:505-513.

Oudbier, J.H., W. Langenberg, E.A.J. Rauws, and C. Bruin-Mosch. 1990. Genotypical variation of *Campylobacter pyloridis* from gastric mucosa. J Clin Microbiol **28**:559-565.

Owen, R.J., S.R Martin, and P Borman. 1985. Rapid urea hydrolysis by gastric *Campylobacters*. Lancet **1(8420)**:111.

Parsonnet, J., G.D. Friedman, D.P. Vandersteen, Y. Chang, J.H. Vogelman, N.Orentreich, and R.K. Sibley. 1991. *Helicobacter pylori* infection in intestinal- and diffuse-type gastric adeno-carcinomas. J Natl Cancer Inst **83**:640-443.

Parsonnet, J., S. Hansen, L. Rodriguez, A.B. Gelb, R.A. Wrnke, E. Jellem, N. Orentreich, J.H.Vogelman, and G.D. Friedman. 1994. *Helicobacter pylori* and primary gastric lymphoma. N Engl J Med **330**:1267-1271.

Paster, B.J., A. Lee, J.G. fox, F.E. Dewhirst, L.A. Tordoff, G.J. Fraser, J.L. O'Rourke, N.S. Taylor, and R. Ferrero. 1991. Phylogeny of *Helicobacter felis* sp. nov., *Helicobacter mustelae*, and related bacteria. Int J Syst Bacteriol **41**(1):31-38.

Phadnis, S.H., T.U. Westablom, and S. Normark. 1993. Molecular cloning of *Helicobacter pylori* DNA: Important differences between *mcrBC* deletion host strains. Mol Microbiol **10**(5):1151.

Phadnis, S.H., D. Ilver, L. Janzon, S. Normark, and T.U. Westablom. 1994. Pathological significance and molecular characterization of the vacuolating toxin gene of *Helicobacter pylori*. Infect Immun **62(5)**:1557-1565.

Raleigh, E.A. and G. Wilson. 1986 *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc.Natl Acad Sci USA **83**:9070-9074.

Ree, H.K., and R.A. Zimmermann. 1990. Organization and expression of the 16S, 23S and 5S ribosomal RNA genes from the archeaebacterium *Thermoplasma acidophilum*. Nucleic Acids Res **18**:4471-4478.

Regensburger, A., W. Ludwig, R. Frank, H. Blocker, and K.H. Schleifer. 1988. Complete nucleotide sequence of a 23S ribosomal RNA gene from *Micrococcus luteus*. Nucleic Acids Res **16**(5):2344.

Romaniuk, P.J., B. Zoltowaska, T.J. Trust,D.J. Lane, G.J. Olsen, N.R. Pace, and D.A. Stahl. 1987. *Campylobacter pyloridis*, the spiral bacterium associated with human gastritis, is not a true *Campylobacter* spp. J Bacteriol **169**:2137-2141.

Sarosiek, J., A. Slomiany, K. Vanhorn, G. Zalesna, and B.L. Slomiany. 1988. Lipolytic activity of *Campylobacter pylori*: Effect of sofalcone. Gastroenterology **94**:A399.

Sarmientos, P. and M. Cashel. 1983. Carbon starvation and growth rate-dependant regulation of the *Escherichia coli* ribosomal RNA promoters: differential control of dual promoters. Proc Natl Acad Sci **80**:1337-1346.

Schmitt, W. and R. Haas. 1994 Genetic analysis of the *Helicobacter pylori* vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein. Mol Microbiol **12(2)**:307-319.

Segal, E., P. Hagblom, H.S. Seifert, and M. So. 1986. Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. Proc Natl Acad Sci USA **83**(7):2177-2181.

Segal, E.D., J. Shon, and L.SD. Tompkins. 1992. Characterization of *Helicobacter pylori* urease mutants. Infect Immun **60**:1883-1889.

Sharma. S.A., M.K.R. Tummuru, G.G. Miller, and M.J. Blaser. 1995. Interleukin-8 response of gastric epithelial cell lines to *Helicobacter pylori* stimulation *in vitro*. Infect Immun **63**(5):1681-1687.

Shull, G.E. and J.B. Lingrel. 1986. Molecular cloning of the rat stomach (H^++K^+)-ATPase. J Biol Chem **261**(36):16788-16791.

Slomiany, B.L., J. Bilski, J. Sarosiek, V.L.N. Murty, B. Dworkin, K. Vanhorn, J. Zielenski, and A. Slomiany. 1987. *Campylobacter pylori* degrades mucin and undermines gastric mucosal integrity. Biochem Biophys Res Com **144**:307-314.

Smith, A.W. B. Chahal, and G.L. French. 1994. The human gastric pathogen *Helicobacter pylori* has a gene encoding an enzyme first classified as a mucinase in *Vibrio cholerae*. Mol Microbiol **13**(1):153-160.

Solioz, M., S. Mathews, and P. Fürst. 1987. Cloning of the K^+-ATPase of *Streptococcus faecalis*. J Biol Chem **262**(15):7358-7362.

Sorokin, A.V., E. Zumstein, V. Azevedo, S.D. Ehrlich, and P. Serror. 1993. The organization of the *Bacillus subtilis* 168 chromosome region between *spo*VA and *ser*A genetic loci, based on sequence data. Mol Microbiol **10**:385-395.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J.Mol Microbiol **98**:503-517.

Spiegelhalder, C., B, Gerstenecker, A. Kersten, E. Schiltz, and M. Kist. 1993. Purification of *Helicobacter pylori* superoxide dismutase and cloning and sequencing of the gene. Infect Immun **61**(12):5315-5325.

Stock, J.B., A.M. Stock, and J. M. Mottonen. 1990. Signal transduction in bacteria. Nature (London) **344**:395-400

Stolte, M. and B. Bethke. 1990. Elimination of *Helicobacter pylori* under treatment with omeprazole. Z Gastroenterol **28**:271-274.

Stucki, U., J. Frey, J. Nicolet, and A.P. Burnes. 1995. Identification of *Campylobacter jejuni* on the basis of a species-specific gene that encodes a membrane protein. J Clin Microbil **33**(4):855-859.

Suerbaum, S., G. Geis, C. Josenhans, and W. Opferkuch. 1992. Biochemical studies of *Helicobacter mustelae* fatty acid composition and flagella. Infect Immun **60**:1695-1698.

Suerbaum, S., C. Josenhans, and A. Labigne. 1993. Cloning and genetic characterization of the *Helicobacter pylori* and *Helicobacter mustelae flaB* flagellin genes and construction of *H. pylori flaA*- and *flaB*- negative mutants by electroporation-mediated allelic exchange. J Bacteriol **175**:3278-3288.

Suerbaum, S., J. Thiberge, I. Kansau, R.L. Ferraro, and A. Labigne. 1994. *Helicobacter pylori hspA-hspB* heat-shock gene cluster: nucleotide sequence, expression, putative function and immunogenicity. Mol Microbiol **14**(5):959-974.

Tait, R.C., T.J. Close, R.C. Lundquist, M. Hagiya, R.L. Rodriguez, and C.I. Kado 1983. Construction and characterization of versatile broad host range DNA cloning system for gram-negative bacteria. Biotechnology **1**:269-275.

Talley, N.J., A.R. Zinsmeister, A. Weaver, E.P. DiMagno, H.A. Carpenter, G.I. Perez-Perez, and M.J. Blaser. 1991. Gastric adenocarcinoma and *Helicobacter pylori* infection. J Natl Cancer Inst **3**:1734-1744.

Taschke, C., M.Q. Klinkert, J. Wolters and R. Herrmann. 1986. Organization of the ribosomal RNA genes in *Mycoplasma hyopneumoniae*: the 5S rRNA gene is

separated from the 16S and 23S rRNA genes. Mol Gen Genet **205**:428-433.

Taylor, D.E., and K. Hiratsuka. 1990. Use of non-radioactive DNA probes for the detection of *Campylobacter jejuni* and *Campylobacter coli* in stool specimens. Mol Cell Probes **4**:261-271.

Taylor, D.E., M. Eaton, N. Chang and S. Salama. 1992a. Construction of a *Helicobacter pylori* genome map and demonstration of diversity at the genome level. J Bacteriol **174**:6800-6806.

Taylor, D.E., M. Eaton, W. Yan, and N. Chang. 1992b. Genome maps of *Campylobacter jejuni* and *Campylobacter coli*. J Bacteriol **174**:2332-2337.

Tee,W., J.R. Lambert, M. Pegorer, and B. Dwyer. 1993. Cytotoxin production by *Helicobacter pylori* more common in peptic ulcer disease. Gastroenterology **104**:A789.

Telford, J.L., P. Ghiara, M Dell'Orco, M. Comanducci, D. Burroni, M. Bugnoli, M.F. Tecce, S. Censini, A. Covacci, Z. Xiang, E. Papini, C. Montecucco, L. Parente, and R. Rappuoli. 1994. Gene structure of the *Helicobacter pylori* cytotoxin and evidence of its key role in gastric disease. J Exp Med **179**:1653-1658.

Triebling, A.T., M.A. Korsten, J.W. Dlugosz, F. Paronetto, and S.S. Lieber. 1991. Severity of *Helicobacter*-induced gastric injury correlates with gastric juice ammonia. Dig Dis Sci **36**:1089-1096.

Trust, T.J., S.M. Logan, C.E. Gustafson, P.J. Romaniuk, N.W. Kim, V.L. Chan, M.A. Ragan, P. Guerry and R.R. Gutell. 1994. Phylogenetic and molecular characterization of a 23S rRNA gene positions the genus *Campylobacter* in the epsilon subdivision of the *Protobacteria* and shows that the presence of transcribed spacers is common in *Campylobacter* spp. J Bacteriol **176**(15):4597-4609.

Tseng, H.C., and C.W. Chen. 1991. A cloned *omp*R-like gene of *Streptomyces lividans* 66 suppresses defective *mel*CI, a putative copper-transfer gene. Mol Microbiol **5**(5):1187-1196.

Tummuru, M.K.R., T.L. Cover, and M.J. Blaser. 1993. Cloning and expression of a high-molecular mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production. Infect Immun **61**:1799-1809.

Tummuru, M.K.R., T.L. Cover, and M.J. Blaser. 1994. Mutaion of the cytotoxin-associated *cag*A gene does not affect the vacuolating cytotoxin activity of *Helicobacter pylori*. Infect Immun **62**)6):2609-2613.

Unge, P., A. Gad, H. Gnarpe, and J. Olsson. 1989. Does omeprazole improve antimicrobiological therapy directed towards gastric *Campylobacter pylori* in patients with antral gastritis. Scan J Gastroenterol **24**(suppl.167):49-54.

van Camp, G., Y. Van de Peer, S. Nicolai, J. Neefs, P. Vandamme and R. deWachter. 1993. Structure of 16S and 23S ribosomal RNA genes in *Campylobacter* species: Phylogenetic analysis of the genus *Campylobacter* and the presence of internal transcribed spacers. Syst Appl Microbiol **16**:361-368.

Vandamme, P., E. Falsen, R. Rossau, B. Hoste, P. Segers, R. Tytgat, and J. de Ley. 1991. Revision of *Campylobacter, Helicobacter,* and *Wolinella* taxonomy: Emendation of generic descriptions and proposal of *Arcobacter* gen. nov. Int J Syst Bacteriol **41**(1):88-103.

Vandamme, P. and J. de Ley. 1991. Proposal for a new family, *Campylobacteraceae*. Int J Syst Bacteriol **41**(3):451-455.

Vandenburghe, A., A. Wassink, P.Raeymaekers, R. De Baere, E. Huysmans, and R. De Wachter. 1985. Nucleotide sequence, secondary structure and evolution of the 5S ribosomal RNA from five bacterial species. Eur J Biochem **149**(3):537-542.

Van den Eynde, H., Y. Van de Peer, J. Perry, and R. De Wachter. 1990. 5S rRNA sequences of representatives of the genera *Chlorobium, Prosthecochloris,* , *Thermomicrobium, Cytophaga, Flavobacterium, Flexibacter,* and *Saprospira* and a discussion of the evolution of eubacteria in general. J Gen Microbiol **136**:11-18.

Vulpe C., B. Levinson, S. Whitney, S. Packman and J. Gitschier. 1993. Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nature Genet **3**:7-13.

Waite-Rees, P.A., C.J. Keating, L.S. Moran, B.,E. Slatko, L.J. Hornstra, and J.S. Benner. 1991. Characterization and expression of the *Escherichia coli* Mrr restriction system. J Bacteriol **173**(6):5207-5219.

Wang, Y. 1988 Double strand dideoxy DNA sequencing. Biotechnology **6**:843-845.

Wang, Y., W.M. Huang, and D.E. Taylor. 1993. Cloning and nucleotide sequence of the *Campylobacter jejuni gyrA* gene and characterization of quinolone resistance mutations. Antimicrob Agents Chemother **37**:457-463.

Warren, J.R. 1983. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet **i**:1273.

Wawrousek, E.F. and N.J. Hansen. 1983. Structure and organization of a cluster of six tRNA genes in the space between tandem ribosomal RNA gene sets in *Bacillus subtilis*. J Biol Chem **258**:291-298.

Weinstock, G.M. 1994. Bacterial genomes: Mapping and stability. ASM News **60**(2):73-78.

Widenhorn, K.A., J. Somers, and W.W. Kay. 1989 Genetic regulation of the tricarboxylate transport operon (*tct*I) of *Salmonella typhimurium*. J Bacteriol **171**:4436-4441.

Woese, C.R. 1987. Bacterial Evolution. Microbiol Rev **51**:221-271.

Wurtzel, E.T., M Chou, and M Inouye. 1982. Osmoregulation of gene expression I: DNA sequence of the *omp*R gene of the *omp*B operon of *Escherichia coli* and characterization of its gene product. J Biol Chem **257**(22):13685-13691.

Wyatt, J.I., B.J. Rathbone, M.F. Dixon, and R.V. Heatly. 1987. *Campylobacter pyloridis* and acid induced gastric metaplasia in the pathogenesis of duodenitis. J Clin Pathol **40**:841-848.

Wyatt, J.I., B.J. Rathbone, G.M. Soballa, T. Shalcrost, R.V. Heatly, A.T.R. Axon, and M.F. Dixon. 1990. Gastric epithelium in the duodenum; its association with *Helicobacter pylori* and inflammation. J Clin Pathol **43**:981-985.

Xiang, Z., S. Censini, P.F. Baeyeli, J.L. Telford, N. Figura, R. Rappuoli, and A. Covacci. 1995. Analysis of expression of CagA and VacA virulence factors in 43 strains of *Helicobacter pylori* reveals that clinical isolates can be divided into two major types and that CagA is not necessary for expression of the vacuolating cytotoxin. Infect Immun **63**(1):94-98.

Yamamoto, J., M. Shimizumu, and K. Yamane. 1991. Molecular cloning and analysis of nucleotide sequences of the *Bacillus subtilis lysA* gene region by using *B.*

subtilis phage vectors and a multi-copy plasmid pUB110. Ag Biol Chem **55**:1615-1626.

Young, R.A. and J.A. Steitz. 1979. Tandem promoters direct *E. coli* ribosomal RNA synthesis. Cell **17**:225-234.

Zhong, P., S.D. Pratt, R.P. Edalji, K.A. Walter, T.F Holzman, A.G. Shivakumar, and L. Katz. 1995. Substrate requirements for ErmC' methyltransferase activity. J Bacteriol **33**(4):4327-4332.

Appendix A

Composition of media and solutions

1.2% denaturing agarose gels (Section F.2)	1.2% (w/v) agarose, 1 X MOPS buffer, 2.2 M formaldehyde
5% stacking gel (Section G)	5% acrylamide (30:0.8(acrylamide:bis-acrylamide)), 50 mM Tris·HCl, pH 6.8
10% separating gel (Section G)	10% acrylamide (30:0.8(acrylamide:bis-acrylamide)), 180 mM Tris·HCl, pH 8.8
1 X MOPS buffer (Section F.2)	40 mM morpholinopropanesulfonic acid, 10 mM NaOAc, 1 mM EDTA, pH 7.0
10 X PM (Section C.4)	0.2% ficoll(Sigma), 0.2% BSA fraction V (Sigma), 0.2% polyvinyl-pyrrolidone (Sigma) in 10 X SSC
20 X SSC	3.0 M NaCl, 3.0 M Na-citrate
Birnboim and Doly solution 1 (Section B.9)	25 mM Tris-HCl pH 7.5, 2% (w/v) glucose, 10 mM EDTA , 50 µg/ml RNase, 10 mg/ml lysozyme
Birnboim and Doly solution 2 (Section B.9)	0.2N NaOH, 1% (w/v) SDS
Birnboim and Doly solution 3 (Section B.9)	3.0 M NaOAc, pH 4.8
blocking buffer (Section C.4)	maleic acid buffer + 1% blocking reagent (BMC)1%
Ca-buffer (Section B.6)	50 mM $CaCl_2$, 10 mM Tris·HCl pH 7.5
CAPS buffer (Section G)	10 mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS), pH 11
DEPC-water (Section F)	200 µl of diethyl pyrocarbonate (DEPC; Aldrich) were added to 100 ml of dH_2O which was incubated at 65°C overnight
denaturing mix (Section F.2)	10% formamide, 2.2 M formaldehyde, 1 X MOPS buffer
denaturing solution (Section B.8)	0.5 M NaOH, 1.5 M NaCl

deproteination solution (Section D.1)	1 mg/ml proteinase K in a solution of 100 mM EDTA, 0.2% (w/v) sodium deoxycholate, 1% (w/v) sarkosyl, 1%(w/v)SDS
destaining solution (Section G)	50% methanol, 10% glacial acetic acid
dNTP mix (Section C.3)	1 mM each of dATP, dCTP, dGTP; 0.65 mM dTTP: 0.35 mM DIG-dUTP; pH 6.5
hybridization solution (Section C.4)	50% formamide, 5 X SSC, 1.0 mM EDTA, 1 x PM, 0.1% SDS, 1% blocking reagent (BMC), 250 µg/ml denatured herring sperm DNA (BMC)
isolation buffer (Section B.1)	0.15 M NaCl, 0.1 M EDTA pH 8.0
LBM-broth (Section B.10)	Luria broth, 10 mM $MgSO_4$, 0.5% (w/v) maltose
loading dye	48% (w/v) sucrose, 0.25% bromophenol blue, 12 mM EDTA, pH 8.0
lysozyme buffer (Section D.1)	10 mg/ml lysozyme in a solution of 100 mM EDTA, 50 mM NaCl, 10 mM Tris·HCl, pH 7.5, 0.2% (w/v) sodium deoxycholate (BDH), 1% (w/v) sarkosyl (BDH)
maleic acid buffer (Section C.4)	100 mM maleic acid, 150 mM NaCl, pH 7.5 (pH adjusted with solid NaOH)
neutralizing solution (Section B.8)	0.5 M Tris·HCl, pH 7.5, 1.5 M NaCl
NT buffer (Section C.3)	25 mM Tris·HCl pH 7.0.5, 5 µM ß-mercaptoethanol,20µM dCTP, 20µM dGTP, 20µM dTTP
NT buffer (10 x) (Section C.3)	250 mM Tris·HCl pH 7.0.5, 50 µM ß-mercaptoethanol, 0.2 mM dCTP, 0.2 mM dGTP, 0.2 mM dTTP
PFGE block washing buffer (Section D.1)	10 mM Tris·HCl, pH 7.0.5, 50 mM EDTA
PFGE running buffer (Section D.3)	0.05 M Tris-base, 0.05 M boric acid, 1 mM EDTA, pH 8.3
PMSF-buffer (Section D.1)	1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma) in PFGE block washing buffer
PPD-buffer (Section C.4)	100 mM Tris·HCl, pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$

protoplast buffer (Section F.1)	15 mM Tris·HCl, pH 8.0, 3.6% sucrose, 8 mM EDTA
SDS-PAGE running buffer (Section G)	25 mM Tris-base, 0.192 M glycine, 0.1% (w/v) SDS
SDS-PAGE running dye (Section G)	2%(w/v) SDS, 5%(w/v) ß-mercaptoethanol, 10%(w/v) glycerol, 62.5 mM Tris·HCl pH 8.3, 0.1%(w/v) bromophenol blue
staining solution (Section G)	0.1% Coomassie blue, 50% methanol
suspension buffer (Section D.1)	0.25 M EDTA, 10 mM Tris·HCl pH 7.5
TAFE running buffer (Section D.3)	5 mM Tris-base, 5 mM boric acid, 0.1 mM EDTA, pH 8.3
TBE-buffer (Section E.3)	0.18M Tris-base, 0.18 M boric acid, 0.2 mM EDTA, pH 8.3
TE	10 mM Tris·HCl, pH 7.5, 1 mM EDTA
TM (Section B.10)	10 mM Tris·HCL pH 7.0.5,10 mM $MgSO_4$
Tris-borate (TB) buffer (0.5 x) (Section C.1)	45 mM Tris base, 45 mM boric acid, 1 mM EDTA, pH 8.3
wash buffer (Section C.4)	0.3% (v/v) Tween 20 in maleic acid buffer

Appendix B

PCR primers used to amplify gene probes from *H. pylori* UA802

The primer sequences are shown from 5'-3'. For each gene probe, the top primer represents the forward (sense) primer and the bottom primer represents the reverse (anti-sense) primer.

Gene	Primer sequence	Reference
*vac*A	-ATGGAAATACAAACACAC- -CGCGCTTTGCACAAATCCAG-	Cover *et al.*, 1994
pfr	-CAACACAAAGGAGATACTATG- TGCACTCAATGACCCTAAAAT-	Pfeifer *et al.*, 1993
*fla*A	-CGGCCCTGAAGCTGTGAGTGAG- -TAACACTTGCAGGATCTATTAC-	Leying *et al.*, 1992
*fla*B	-CAGCGCTCGTAGAAATGCGCACGG- -GTTTATACCTATTAATGAATG-	Suerbaum *et al.*, 1993
*cag*A	-GATAACGCTGTCGCTTCATACG- -ACCTGCTATGACTAAGCCACTG-	Tummuru *et al.*, 1993
*lpp*20	-TTAGGGTAGAGTGTGGTAGCT- -ACCTGCTATGACTAAGCCACTG-	Kostrzynska *et al.*, 1994
*gyr*B	-ARMGNCCNGGNATGTAYATHGGN- -NGGDATNCCRCKNCCRTCRTC-	Huang, 1992
*kat*A	-GTAGTGCATGCTAAAGGAAGC- -GGTAACATGAGCTAAAGACTC-	Newell *et al.*, 1991
*hpa*A	-GATAGAACGATGAAACAAATG- -GTCTAACTCACTCAACTCCAT-	Evans *et al.*, 1993

R- A/G Y- T/C M- A/C K- G/T S-G/C W- A/T
H- A/C/T B- C/G/T V- A/C/G D- A/G/T N- A/C/G/T

Appendix C
Comparisons of rRNA gene sequences from *Bacillus subtilis*

Sequences shown below, show comparisons of the 23S and 5S genes from six different rRNA operons from *B. subtilis*, as well as the flanking regions and intergenic spacer. The 5'-flanking region and first 107 bp of the 23S gene from the *rrn*I operon were not available.

References: rrnI - Warousek and Hansen, 1983
rrnO, rrnA, rrnJ and rrn W - Ogasawara *et al.*, 1994
rrnB - Green *et al.*, 1985

The start and end sites of the 23S and 5S genes are indicated by shaded boxes and are labelled.

```
      1                                                                    70
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  AGACCTTGGG TCTTATAAAC AGAACGTTCC CTGTCTTGTT TAGTTTTGAA GGAACTTTGT TCCTTGAATA
rrnA  GCCTTGTACA CACCGCCCGT CACACCACGA GAGTTTGTAA CACCCGAAGT CGGTGAGGTA ACCTTTTAGG
rrnJ  GCCTTGTACA CACCGCCCGT CACACCACGA GAGTTTGTAA CACCCGAAGT CGGTGAGGTA ACCTTTTAGG
rrnW  GCCTTGTACA CACCGCCCGT CACACCACGA GAGTTTGTAA CACCCGAAGT CGGTGAGGTA ACCTTTTAGG
rrnB  AGACCTTGGG TCTTATAAAC AGAACGTTCC CTGTCTTGTT TAGTTTTGAA GGAACTTTGT TCCTTGAATA

      71                                                                   140
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  AGTTAAGATG GGCCTGTAGC TCAGCTGGTT AGAGCGCACG CCTGATAAGC GTGAGGTCGG TGGTTCGAGT
rrnA  AGCCAGCCGC CGAAGGTGGG ACAGATGATT GGGGTGAAGT CGTAACAAGG TAGCCGTATC GGAAGGTGCG
rrnJ  AGCCAGCCGC CGAAGGTGGG ACAGATGATT GGGGTGAAGT CGTAACAAGG TAGCCGTATC GGAAGGTGCG
rrnW  AGCCAGCCGC CGAAGGTGGG ACAGATGATT GGGGTGAAGT CGTAACAAGG TAGCCGTATC GGAAGGTGCG
rrnB  AGTTAAGATG GGCCTGTAGC TCAGCTGGTT AGAGCGCACG CCTGATAAGC GTGAGGTCGG TGGTTCGAGT

      141                                                                  210
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  CCACTCAGGC CCACCATCTT ATATAACGGG GCCTTAGCTC AGCTGGGAGA GCGCCTGCTT TGCACG...C
rrnA  GCTGGATCAC CTCCTTTCTA AGGATATTTT ACGGAATATA AGACCTTGGG TCTTATAAAC AGAACGTTCC
rrnJ  GCTGGATCAC CTCCTTTCTA AGGATATTTT ACGGAATATA AGACCTTGGG TCTTATAAAC AGAACGTTCC
rrnW  GCTGGATCAC CTCCTTTCTA AGGATATTAT ACGGAATATA AGACCCAAGG TCTTATAAAC AGAACGTTCC
rrnB  CCACTCAGGC CCACCATCTT ATATAACGGG GCCTTAGCTC AGCTGGGAGA GCGCCTGCTT TGCACG...C

      211                                                                  280
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  AGGAGGTCAG CGGTTCGATC CCGCTAGGCT CCACCAACGT GTTCTTTGAA AACTAGATAA CAGTAGACAT
rrnA  CTGTCTTGTT TAGTTTTGAA GGATCATTCC TTCGAAACGT GTTCTTTGAA AACTAGATAA CAGTAGACAT
rrnJ  CTGTCTTGTT TAGTTTTGAA GGATCATTCC TTCGAAACGT GTTCTTTGAA AACTAGATAA CAGTAGACAT
rrnW  CTGTCTTGTT TAGTTTTGAA GGATCATTCC TTCGAAACGT GTTCTTTGAA AACTAGATAA CAGTAGACAT
rrnB  AGGAGGTCAG CGGTTCGATC CCGCTAGGCT CCACCAACGT GTTCTTTGAA AACTAGATAA CAGTAGACAT

      281                                    23S gene starts               350
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  CACATTCAAT TAGTAACACA AGATATCACA TAGTGATTCT TTTTAACGGT TAAGTTAGAA AGGGCGCACG
rrnA  CACATTCA.. .AGTAATACA AGATATCACA TAGTGATTCT TTTTAACGGT TAAGTTAGAA AGGGCGCACG
rrnJ  CACATTCA.. .AGTAATACA AGATATCACA TAGTGATTCT TTTTAACGGT TAAGTTAGAA AGGGCGCACG
rrnW  CACATTCAAT TAGTAACACA AGATATCACA TAGTGATTCT TTTTAACGGT TAAGTTAGAA AGGGCGCACG
rrnB  CACATTCAAT TAGTAACACA AGATATCACA TAGTGATTCT TTTTAACGGT TAAGTTAGAA AGGGCGCACG

      351                                                                  420
rrnI  .......... .......... .......... .......... .......... .......... ..........
rrnO  GTGGATGCCT TGGCACTAGG AGCCGATGAA GGACGGGACG AACACCGATA TGCTTCGGGG AGCTGTAAGC
rrnA  GTGGATGCCT TGGCACTAGG AGCCGATGAA GGACGGGACG AACACCGATA TGCTTCGGGG AGCTGTAAGC
rrnJ  GTGGATGCCT TGGCACTAGG AGCCGATGAA GGACGGGACG AACACCGATA TGCTTCGGGG AGCTGTAAGC
rrnW  GTGGATGCCT TGGCACTAGG AGCCGATGAA GGACGGGACG AACACCGATA TGCTTCGGGG AGCTGTAAGC
rrnB  GTGGATGCCT TGGCACTAGG AGCCGATGAA GGACGGGACG AACACCGATA TGCTTCGGGG AGCTGTAAGC

      421                                                                  490
rrnI  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT
```

```
rrnO  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT
rrnA  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT
rrnJ  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT
rrnW  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT
rrnB  AAGCTTTGAT CCGGAGATTT CCGAATGGGG AAACCCACCA CTCGTAATGG AGTGGTATCC ATATCTGAAT

          491                                                                  560
rrnI  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGAGGAAGAG AAAGCAAATG
rrnO  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGAGGAAGAG AAAGCAAATG
rrnA  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGAGGAAGAG AAAGCAAATG
rrnJ  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGAGGAAGAG AAAGCAAATG
rrnW  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGA.GAAGAG AAAGCAAATG
rrnB  TCATAGGATA TGAGAAGGCA GACCCGGGGA ACTGAAACAT CTAAGTACCC GGAGGAAGAG AAAGCAAATG

          561                                                                  630
rrnI  CGATTCCCTG AGTAGCGGCG AGCGAAACGG GATCAGCCCA AACCAAGAGG CTTGCCTCTT GGGGTTGTAG
rrnO  CGATTCCCTG AGTAGCGGCG AGCGAAACGG GATCAGCCCA AACCAAGAGG CTTGCCTCTT GGGGTTGTAG
rrnA  CGATTCCCTG AGTAGCGGCG AGCGAAACGG GATTAGCCCA AACCAAGAGG CTTGCCTCTT GGGGTTGTAG
rrnJ  CGATTCCCTG AGTAGCGGCG AGCGAAACGG GATTAGCCCA AACCAAGAGG CTTGCCTC.T GGGGTTGTAG
rrnW  CGATTCCCTG AGTAGCGGCG ACGAACACGG GATCAGCCCA AACCAAGAGG CTTGCCTC.T GTGGTTGTAG
rrnB  CGATTCCCTG AGTAGCGGCG AGCGAAACGG GATTAGCCCA AACCAAGAGG CTTGCCTCTT GGGGTTGTAG

          631                                                                  700
rrnI  GACACTCTGT ACGGAGTTAC AAAGGAACGA GGTAGATGAA GAGGTCTGGA AA.GGCCCGC CATAGGAGGT
rrnO  GACACTCTGT ACGGAGTTAC AAAGGAACGA GGTAGATGAA GAGGTCTGGA AA.GGCCCGC CATAGGAGGT
rrnA  GACACTCTGT ACGGAGTTAC AAAGGAACGA GGTAGATGAA GAGGTCTGGA AA.GGCCCGC CATAGGAGGT
rrnJ  GACACTCTGT ACGGAGTTAC AAAGGAACGA GGTAGATGAA GAGGTCTGGA AA.GGCCCGC CATAGGAGGT
rrnW  GACACTCTGT ACGGAGTTAC AAAAGAACGA GGTAGATGAA GAGGTCTGGA AAGGGCCCGC CATAGGAGGT
rrnB  GACACTCTGT ACGGAGTTAC AAAGGAACGA GGTAGATGAA GAGGTCTGGA AA.GGCCCGC CATAGGAGGT

          701                                                                  770
rrnI  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC
rrnO  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC
rrnA  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC
rrnJ  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC
rrnW  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC
rrnB  AACAGCCCTG TAGTCAAAAC TTCGTTCTCT CCTGAGTGGA TCCTGAGTAC GGCGGAACAC GTGAAATTCC

          771                                                                  840
rrnI  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG
rrnO  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG
rrnA  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG
rrnJ  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG
rrnW  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG
rrnB  GTCGGAATCC GGGAGGACCA TCTCCCAAGG CTAAATACTC CCTAGTGACC GATAGTGAAC CAGTACCGTG

          841                                                                  910
rrnI  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA
rrnO  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA
rrnA  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA
rrnJ  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA
rrnW  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA
rrnB  AGGGAAAGGT GAAAAGCACC CCGGAAGGGG AGTGAAAGAG ATCCTGAAAC CGTGTGCCTA CAAGTAGTCA

          911                                                                  980
rrnI  GAGCCCGTTA ACGGGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCTCG TGCAAGGTTA
rrnO  GAGCCCGTTA ACGGGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCCCG TGCAAGGTTA
rrnA  GAGCCCGTTA ACGGGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCCCG TGCAAGGTTA
rrnJ  GAGCCCGTTA ACGGGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCTCG TGCAAGGTTA
rrnW  GAGCCCGTTA AC.GGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCCCG TGCAAGGTTA
rrnB  GAGCCCGTTA ACGGGTGATG GCGTGCCTTT TGTAGAATGA ACCGGCGAGT TACGATCCCG TGCAAGGTTA

          981                                                                 1050
rrnI  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA
rrnO  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA
rrnA  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA
```

```
rrnJ  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA
rrnW  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA
rrnB  AGCAGAAGAT GCGGAGCCGC AGCGAAAGCG AGTCTGAATA GGGCGCATGA GTACGTGGTC GTAGACCCGA

      1051                                                                 1120
rrnI  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG
rrnO  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG
rrnA  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG
rrnJ  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG
rrnW  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG
rrnB  AACCAGGTGA TCTACCCATG TCCAGGGTGA AGTTCAGGTA ACACTGAATG GAGGCCCGAA CCCACGCACG

      1121                                                                 1190
rrnI  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAAC CTGGAGATAG CTGGTTCTCT
rrnO  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAAC CTGGAGATAG CTGGTTCTCT
rrnA  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAAC CTGGAGATAG CTGGTTCTCT
rrnJ  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAGC CTGGAGATAG CTGGTTCTCT
rrnW  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAAC CTGGAGATAG CTGGTTCTCT
rrnB  TTGAAAAGTG CGGGGATGAG GTGTGGGTAG GGGTGAAATG CCAATCGAAC CTGGAGATAG CTGGTTCTCT

      1191                                                                 1260
rrnI  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC
rrnO  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC
rrnA  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC
rrnJ  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC
rrnW  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC
rrnB  CCGAAATAGC TTTAGGGCTA GCCTCAAGGT AAGAGTCTTG GAGGTAGAGC ACTGATTGGA CTAGGGGCCC

      1261                                                                 1330
rrnI  CTACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG
rrnO  CTACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG
rrnA  CTACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG
rrnJ  CTACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG
rrnW  TCACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG
rrnB  CTACCGGGTT ACCGAATTCA GTCAAACTCC GAATGCCAAT GACTTATCCT TGGGAGTCAG ACTGCGAGTG

      1331                                                                 1400
rrnI  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA
rrnO  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA
rrnA  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA
rrnJ  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA
rrnW  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA
rrnB  ATAAGATCCG TAGTCGAAAG GGAAACAGCC CAGACCGCCA GCTAAGGTCC CAAAGTATAC GTTAAGTGGA

      1401                                                                 1470
rrnI  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA
rrnO  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA
rrnA  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA
rrnJ  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA
rrnW  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA
rrnB  AAAGGATGTG GAGTTGCTTA GACAACCAGG ATGTTGGCTT AGAAGCAGCC ACCATTTAAA GAGTGCGTAA

      1471                                                                 1540
rrnI  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT
rrnO  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT
rrnA  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT
rrnJ  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT
rrnW  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT
rrnB  TAGCTCACTG GTCGAGTGAC TCTGCGCCGA AAATGTACCG GGGCTAAACG TATCACCGAA GCTGCGGACT

      1541                                                                 1610
rrnI  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGAGCGCT
rrnO  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGAGCGCT
```

```
rrnA  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGAGCGCT
rrnJ  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGAGCGCT
rrnW  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGACGGCT
rrnB  GTTCTTCGAA CAGTGGTAGG AGAGCGTTCT AAGGGCTGTG AAGCCAGACC GGAAGGACTG GTGGAGCGCT

      1611                                                                  1680
rrnI  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAGAGGGGTG AGAATCCCCT CCACCGAATG CCTAAGGTTT
rrnO  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAGAGGGGTG AGAATCCCCT CCACCGAATG CCTAAGGTTT
rrnA  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAGAGGGGTG AGAATCCCCT CCACCGAATG CCTAAGGTTT
rrnJ  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAGAGGGGTG AGAATCCCCT CCACCGAATG CCTAAGGTTT
rrnW  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAAGAGGGGT GAGAATCCCT CCACCGAATG CCTAAGGGTT
rrnB  TAGAAGTGAG AATGCCGGTA TGAGTAGCGA AAGAGGGGTG AGAATCCCCT CCACCGAATG CCTAAGGTTT

      1681                                                                  1750
rrnI  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC
rrnO  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC
rrnA  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC
rrnJ  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC
rrnW  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC
rrnB  CCTGAGGAAG GCTCGTCCGC TCAGGGTTAG TCGGGACCTA AGCCGAGGCC GAAAGGCGTA GGCGATGGAC


      1751                                                                  1820
rrnI  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG ACGCAGGAGG ATAGGGTAAG
rrnO  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG ACGCAGGAGG ATAGGGTAAG
rrnA  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG ACGCAGGAGG ATAGGGTAAG
rrnJ  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG ACGCAGGAGG ATAGGGTAAG
rrnW  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG TCGCAGGAGG ATAGGGTAAG
rrnB  AACAGGTTGA TATTCCTGTA CCACCTCCTC ACCATTTGAG CAATGGGGGG ACGCAGGAGG ATAGGGTAAG

      1821                                                                  1890
rrnI  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA
rrnO  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA
rrnA  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA
rrnJ  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA
rrnW  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA
rrnB  CGCGGTATTG GATATCCGCG TCCAAGCAGT TAGGCTGGGA AATAGGCAAA TCCGTTTCCC ATAAGGCTGA

      1891                                                                  1960
rrnI  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG
rrnO  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG
rrnA  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG
rrnJ  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG
rrnW  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG
rrnB  GCTGTGATGG CGAGCGAAAT ATAGTAGCGA AGTTCCTGAT TCCACACTGC CAAGAAAAGC CTCTAGCGAG

      1961                                                                  2030
rrnI  GTGAGAGGTG CCCGTACCGC AAACCGACAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT
rrnO  GTGAGAGGTG CCCGTACCGC AAACCGACAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT
rrnA  GTGAGAGGTG CCCGTACCGC AAACCGACAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT
rrnJ  GTGAGAGGTG CCCGTACCGC AAACCGACAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT
rrnW  GTGAGAGGTG CCCGTACCGC AAACCGTCAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT
rrnB  GTGAGAGGTG CCCGTACCGC AAACCGACAC AGGTAGGCGA GGAGAGAATC CTAAGGTGAT CGAGAGAACT

      2031                                                                  2100
rrnI  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC
rrnO  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC
rrnA  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC
rrnJ  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC
rrnW  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC
rrnB  CTCGTTAAGG AACTCGGCAA AATGACCCCG TAACTTCGGG AGAAGGGGTG CTCTGTTAGG GTGCAAGCCC

      2001                                                                  2170
rrnI  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC
rrnO  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC
rrnA  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC
```

```
rrnJ  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC
rrnW  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC
rrnB  GAGAGAGCCG CAGTGAATAG GCCCAGGCGA CTGTTTAGCA AAAACACAGG TCTCTGCGAA GCCGTAAGGC

      2171                                                                  2240
rrnI  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG
rrnO  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG
rrnA  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG
rrnJ  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG
rrnW  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG
rrnB  GAAGTATAGG GGCTGACGCC TGCCCGGTGC TGGAAGGTTA AGAGGAGCGC TTAGCGTAAG CGAAGGTGCG

      2241                                                                  2310
rrnI  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA
rrnO  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA
rrnA  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA
rrnJ  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA
rrnW  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA
rrnB  AATTGAAGCC CCAGTAAACG GCGGCCGTAA CTATAACGGT CCTAAGGTAG CGAAATTCCT TGTCGGGTAA

      2311                                                                  2380
rrnI  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCACTGT CTCAACGAGA GACTCGGTGA AATTATAGTA
rrnO  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCACTGT CTCAACGAGA GACTCGGTGA AATTATAGTA
rrnA  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCACTGT CTCAACGAGA GACTCGGTGA AATTATAGTA
rrnJ  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCACTGT CTCAACGAGA GACTCGGTGA AATTATAGTA
rrnW  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCGCTGT CTCAACGAGA GACTCGGTGA AATTATAGTA
rrnB  GTTCCGACCC GCACGAAAGG CGCAACGATC TGGGCACTGT CTCAACGAGA GACTCGGTGA AATTATAGTA



      2381                                                                  2450
rrnI  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG
rrnO  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG
rrnA  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG
rrnJ  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG
rrnW  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG
rrnB  CCTGTGAAGA TGCAGGTTAC CCGCGACAGG ACGGAAAGAC CCCGTGGAGC TTTACTGCAG CCTGATATTG

      2451                                                                  2520
rrnI  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTCG GTGGAGGCAT
rrnO  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTTG GTGGAGGCAT
rrnA  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTCG GTGGAGGCAT
rrnJ  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTCG GTGGAGGCAT
rrnW  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTCG GTGGAGGCAT
rrnB  AATGTTGGTA CAGCTTGTAC AGGATAGGTA GGAGCCTTGG AAACCGGAGC GCCAGCTTCG GTGGAGGCAT

      2521                                                                  2590
rrnI  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCG CCGCCCTTAT CGGGCGGGGA GACAGTGTCA
rrnO  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCG CCGCCCTTAT CGGGCGGGGA GACAGTGTCA
rrnA  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCG CCGCCCTTAT CGGGCGGGGA GACAGTGTCA
rrnJ  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCG CCGCCCTTAT CGGGCGGGGA GACAGTGTCA
rrnW  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCG CCGCCCTTAT CGGGCGGGGA GACAGTGTCA
rrnB  CGGTGGGATA CTACCCTGGC TGTATTGACC TTCTAACCCC CCGCCCTTAT CGGGCGGGGA GACAGTGTCA

      2591                                                                  2660
rrnI  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CTAAAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG
rrnO  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CTAAAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG
rrnA  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CTAAAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG
rrnJ  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CT.AAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG
rrnW  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CTAAAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG
rrnB  GGTGGGCAGT TTGACTGGGG CGGTCGCCTC CTAAAAGGTA ACGGAGGCGC CCAAAGGTTC CCTCAGAATG

      2661                                                                  2730
rrnI  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA
rrnO  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA
rrnA  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA
```

```
rrnJ  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA
rrnW  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA
rrnB  GTTGGAAATC ATTCGCAGAG TGTAAAGGCA CAAGGGAGCT TGACTGCGAG ACCTACAAGT CGAGCAGGGA

      2731                                                                    2800
rrnI  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC
rrnO  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC
rrnA  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC
rrnJ  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC
rrnW  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC
rrnB  CGAAAGTCGG GCTTAGTGAT CCGGTGGTTC CGCATGGAAG GGCCATCGCT CAACGGATAA AAGCTACCCC

      2801                                                                    2870
rrnI  GGGGATAACA GGCTTATCTC CCCCAAGAG. TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC
rrnO  GGGGATAACA GGCTTATCTC CCCCAAGAG. TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC
rrnA  GGGGATAACA GGCTTATCTC CCCCAAGAG. TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC
rrnJ  GGGGATAACA GGCTTATCTC CCCCAAGAG. TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC
rrnW  GGGGATAACA GGCTTATCTC CCCCAAGAGC TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC
rrnB  GGGGATAACA GGCTTATCTC CCCCAAGAG. TCCACATCGA CGGGGAGGTT TGGCACCTCG ATGTCGGCTC

      2871                                                                    2940
rrnI  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG
rrnO  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG
rrnA  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG
rrnJ  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG
rrnW  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG
rrnB  ATCGCATCCT GGGGCTGTAG TCGGTCCCAA GGGTTGGGCT GTTCGCCCAT TAAAGCGGTA CGCGAGCTGG

      2941                                                                    3010
rrnI  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCAGGA AATTTGAGAG GAGCTGTCCT
rrnO  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCAGGA AATTTGAGAG GAGCTGTCCT
rrnA  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCAGGA AATTTGAGAG GAGCTGTCCT
rrnJ  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCAGGA AATTTGAGAG GAGCTGTCCT
rrnW  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCTGGA AATTTGAGAG GAGCTGTCCT
rrnB  GTTCAGAACG TCGTGAGACA GTTCGGTCCC TATCCGTCGC GGGCGCAGGA AATTTGAGAG GAGCTGTCCT

      3011                                                                    3080
rrnI  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG
rrnO  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG
rrnA  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG
rrnJ  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG
rrnW  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG
rrnB  TAGTACGAGA GGACCGGGAT GGACGCACCG CTGGTGTACC AGTTGTTCTG CCAAGGGCAT CGCTGGGTAG

      3051                                                                    3150
rrnI  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT
rrnO  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT
rrnA  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT
rrnJ  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT
rrnW  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT
rrnB  CTATGTGCGG ACGGGATAAG TGCTGAAAGC ATCTAAGCAT GAAGCCCCCC TCAAGATGAG ATTTCCCATT

      3151                                                                    3220
rrnI  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
rrnO  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
rrnA  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
rrnJ  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
rrnW  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
rrnB  CCGCAAGGAA GTAAGATCCC TGAAAGATGA TCAGGTTGAT AGGTCTGAGG TGGAAGTGTG GCGACACATG
                                        23S gene ends
      3221                                                                    3290
rrnI  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCATATT TTTGAATGAT GTCACACCTG TTATCTAGTT
rrnO  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCATATT TTTGAATGAT GTCACATCTG TTATCTAGTT
rrnA  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCATATT TTTGAATGAT GTCACATCTG TTATCTAGTT
```

```
rrnJ  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCATATT TTTGAATGAT GTCACATCTG TTATCTAGTT
rrnW  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCATATT TTTGAATGAT GTCACACCTG TTATCTAGTT
rrnB  GAGCTGACAG ATACTAATCG ATCGAGGACT TAACCACA.T TTTGAATGAT GTCACACCTG TTATCTAGTT

      3291                                                                  3360
rrnI  TTGAGAGAAC AC..... ..... ..... ..... ..... .....
rrnO  TTGAGAGAAC AC..... ..... ..... ..... ..... .....
rrnA  TTGAGAGAAC AC..... ..... ..... ..... ..... .....
rrnJ  TTGAGAGAAC AC..... ..... ..... ..... ..... .....
rrnW  TTGAGAGAAC AC..... ..... ..... ..... ..... .....
rrnB  TTGAGAGAAC ACCTCTCTAA AGGCGGAAGG TAAGGAAACT CCGCTAAGGG CTCTCACATC CTGTGAGAAA
                              5S gene begins
      3361                    ▼                                             3430
rrnI  ..... ..... ..... TCTCAATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC
rrnO  ..... ..... ..... TCTCAATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC
rrnA  ..... ..... ..... TCTCAATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC
rrnJ  ..... ..... ..... TCTCAATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC
rrnW  ..... ..... ..... TCTCAATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC
rrnB  CGCCCAGTAC CTTCATCCTG AAGGCATTTG TTTGGTGGCG ATAGCGAAGA GGTCACACCC GTTCCCATAC

      3431                                                                  3500
rrnI  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTAGTCG GGGGTTTCCC CCTGTGAGAG TAGGACGCCG
rrnO  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTAGTCG GGGGTTTCCC CCTGTGAGAG TAGGACGCCG
rrnA  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTAGTCG GGGGTTTCCC CCTGTGAGAG TAGGACGCCG
rrnJ  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTAGTCG GGGGTTTCCC CCTGTGAGAG TAGGACGCCG
rrnW  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTAGTCG GGGGTTTCCC CCTGTGAGAG TAGGACGCCG
rrnB  CGAACACGGA AGTTAAGCTC TTCAGCGCCG ATGGTA.... ..... ..... .....
      5S ends
      3501                                    3544
rrnI  CCAAGCAAGC TTAAACCCAG CTCAATGAGC TGGGTTTTTT GTT.
rrnO  CCAAGCAAGC TTAAACCCAG CTCAATGAGC TGGGTTTTTT GTA.
rrnA  CCAAGCAATT GCACGTTAGT GCAATTATGG AGGATTAGCT CAG.
rrnJ  CCAAGCAGAT TTATTAATTC TACCATGTTT GTTTATGAAG CTT.
rrnW  CCAAGCAATT GCACG.TTAG TGCAATATGG AGGATTAGCT CAGC
rrnB  ..... ..... ..... ..... ....
```